

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 05:48:35 ; Search time 370 Seconds
(without alignments)
6337.256 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcgacggtgttc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/PCRTUS COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.8	7.9	31491	3	US-09-360-186-1
2	113.8	7.9	31491	4	US-09-864-680A-1
3	101.8	7.1	612	4	US-09-902-540-1357
4	90.2	6.3	59519	4	US-09-949-016-13504
5	88.8	6.2	55886	4	US-09-949-016-15129
6	88.4	6.2	19124	2	US-08-487-826B-13
7	88.4	6.2	205044	4	US-09-949-016-15851
8	88.4	6.2	205044	4	US-09-949-016-15852
9	88.4	6.2	205044	4	US-09-949-016-15853
10	88.4	6.2	223471	4	US-09-949-016-12387
11	88.4	6.2	223471	4	US-09-949-016-12724
12	88.4	6.2	223471	4	US-09-949-016-12725
13	87.4	6.1	61178	4	US-09-949-016-17369
14	85.8	6.0	187169	4	US-09-949-016-12776
15	85.8	6.0	191569	4	US-09-949-016-15940
16	84.4	5.9	1039	4	US-09-902-540-1280
17	84.2	5.9	263693	4	US-09-949-016-12386
18	84.2	5.9	263694	4	US-09-949-016-16915
19	84	5.9	20674	3	US-09-641-638-651
20	84	5.9	20674	4	US-10-170-097-651
21	83.8	5.8	147382	4	US-09-949-016-14624
22	83.6	5.8	601	4	US-09-949-016-30530
23	83.6	5.8	601	4	US-09-949-016-30531
24	83.6	5.8	601	4	US-09-949-016-37149
25	83.6	5.8	601	4	US-09-949-016-37150
26	83.6	5.8	601	4	US-09-949-016-37163
27	83.6	5.8	601	4	US-09-949-016-37164

28	83.6	5.8	601	4	US-09-949-016-145867	Sequence 145867, A
29	83.6	5.8	601	4	US-09-949-016-145868	Sequence 145868, A
30	83.6	5.8	601	4	US-09-949-016-146135	Sequence 146135, A
31	83.6	5.8	601	4	US-09-949-016-146136	Sequence 146136, A
32	83.6	5.8	601	4	US-09-949-016-146403	Sequence 146403, A
33	83.6	5.8	601	4	US-09-949-016-146404	Sequence 146404, A
34	83.4	5.8	615	3	US-08-998-416-186	Sequence 186, App
35	83.4	5.8	837	3	US-08-998-416-288	Sequence 288, App
36	82.6	5.8	194388	4	US-09-949-016-12699	Sequence 12699, A
37	81.6	5.7	18773	4	US-09-949-016-14164	Sequence 14164, A
38	81.6	5.7	95255	4	US-09-949-016-17067	Sequence 17067, A
39	81	5.7	18773	4	US-09-949-016-14164	Sequence 14164, A
40	80.8	5.6	636	3	US-08-998-416-1137	Sequence 1137, App
41	80.6	5.6	658	3	US-08-998-416-595	Sequence 595, App
42	80.2	5.6	29717	4	US-09-949-016-16284	Sequence 16284, A
43	80.2	5.6	60376	4	US-09-949-016-12423	Sequence 12423, A
44	79.8	5.6	134987	4	US-09-949-016-15348	Sequence 15348, A
45	79.8	5.6	134987	4	US-09-949-016-15349	Sequence 15349, A

ALIGNMENTS

RESULT 1
US-09-360-186-1
; Sequence 1, Application US/09360186
; Patent No. 6262343
; GENERAL INFORMATION:
; APPLICANT: Staskiewicz, et al.
; TITLE OF INVENTION: B2 Resistance Gene
; FILE REFERENCE: 50687
; CURRENT APPLICATION NUMBER: US/09/360,186
; EARLIER FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: 60/093,957
; EARLIER FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 31491
; TYPE: DNA
; ORGANISM: Capsicum annuum
; US-09-360-186-1

Query Match	7.9%	Score 113.8;	DB 3;	Length 31491;
Best Local Similarity	59.5%	Pred. No. 2.9e-14;		
Matches	210;	Conservative	0;	Mismatches 142; Indels 1; Gaps 1;
Qy	38	CTTGATATCATGATTTTCAGATTGTCATGAAGACTTCTATCTATCAGAACGCTGCAG	97	
Db	22086	CTCTAGTTCTTCTTCGAAACATCGGCATAGGACATCTAACGACTTTTGAGCAGTCTTGAC	22145	
Qy	98	AGGATCCCAATTTAGTCTAAATTTATCTTCAGTCTCGAAACCAACTCAGGACCCAAAC	157	
Db	22146	TGCGTCTCTAATGATTTTCACTTCTTCATCGCCCTATGAAACAGATTAGGCCATACAA	22204	
Qy	158	CCGTGCTCACCAACTCAGTCTAATATAACAGAGTATGACATTAATGACATATAGC	217	
Db	22205	CTGAGTCTACCCACTTCTATACCATCTATCGGAGACCTATATCTCTCAATCAAAAGC	22264	
Qy	218	CTGTAAGTGCATCTAGATGCGAGATTGGAACCTGTTATTGAGGGGAACCTCACTAA	277	
Db	22265	CTCAAAAGGAGCCATCTTGATGCTGCGATGTTATTTATTTGTAAGCAATTCACCGAG	22324	
Qy	278	CGGTAAATTAATCTTCAACTTCTAGTAATAATCATACTAGCTCCAAATCGTATCTCT	337	
Db	22325	TGGCAGGTGATCTACCCAACTACTCTTTGAAATCAATTCAGTCCCTTAACATATCTTC	22384	
Qy	338	TAGTATATGAATCACCTTCTCAAAATTCACCATCGCTGAGGATGGAATGCGAG	390	
Db	22385	GAGGTTCTGAATGGTAGCCTCAGCTTCTCCATCGTCTGAGGTGGAACCTG	22437	

RESULT 2

US-09-864-680A-1
; Sequence 1, Application US/09864680A
; Patent No. 6762285
; GENERAL INFORMATION:
; APPLICANT: Staskiewicz, Brian J
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Tai, Thomas H
; TITLE OF INVENTION: B2 RESISTANCE GENE
; FILE REFERENCE: 42250/234021 (5830-4A)
; CURRENT APPLICATION NUMBER: US/09/864,680A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/360,186
; PRIOR FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 60/093,957
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 31491
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-864-680A-1

Query Match 7.9%; Score 113.8; DB 4; Length 31491;
Best Local Similarity 59.5%; Pred. No. 2.9e-14;
Matches 210; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 38 CTGTGATCTATGATTTTCAGATTGTCATGAAGACTTCTATCTATCAGAGACGCCCTGCAG 97
DB 22086 CTCTAGTCTCTCTCCGAAACATCGGCATAGGACATCTAAGACTTTTGAGCAGCTTGAC 22145
QY 98 AGGATCCCAATAGTCTAAATATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
DB 22146 TCG-TCCTATGATTTTCACTTCTCCATCGCTTATGAACAAGATTAGGCCCATACAA 22204
QY 158 CCGTCGCTACCCCACTCAGTCTAATATACAGATGACACTTATGACCATATAGAGC 217
DB 22205 CTGAGTCTACCCCACTTATACCATCTCTATCGGAGACCTATATCTCTCAATACAAAGC 22264
QY 218 CTGCTAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATGTCGGAACCTCAACTAA 277
DB 22265 CTCAAAAGGAGCCATCTGTGATGTCGATGTTATTTATGTAAGCGAATTCACCCAG 22324
QY 278 CGGTAAAAATCCTCTCAACTACCTTAGTAAATACATAGTCTCAAAATCGTATCCTC 337
DB 22325 TGGCAGGTGATCTACCCCACTCTTGAATCAATTCAGCATGCCCTAACATATCTC 22384
QY 338 TAGTATATGAATCACCTTCTCAAAATGACCATCGGTCTGAGGATGGAAATGCAG 390
DB 22385 GAGGGTCTGAATGGTACGCTCAGCTTGTCCATCGTCTGAGGGTGGAAACCTG 22437

RESULT 3
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 7.1%; Score 101.8; DB 4; Length 612;
Best Local Similarity 48.1%; Pred. No. 3.7e-12;
Matches 286; Conservative 0; Mismatches 309; Indels 0; Gaps 0;

QY 502 TTTTATTATTACCTCCTCGTAGATTGTAATAACTATTGCAATTTATCTCAATTCATTA 561
DB 611 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 552
QY 562 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 621
DB 551 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 492
QY 622 AGAATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 681
DB 491 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 432
QY 682 TAAAGAAGAAATTCGAATTGCGTTTCTTAATTAATTAGTTTAAATCTATATACTAATA 741
DB 431 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 372
QY 742 AATTCTGATCGGATTAGTGTGTCAAAGTCAAGTCACATGAATTTGTGTGGAGAAAAA 801
DB 371 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 312
QY 802 ATAAAAATTAAACACATTTTCGATTAAATTATATATATATATATATATATATATATA 861
DB 311 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 252
QY 862 TTATTAAATGTTGTCGAATAATAATTTTAAATTAATAATTTTAAATTTTAAATTTTAA 921
DB 251 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 192
QY 922 ATCAATTAATTTAAATCTTATTACCATAAATAATAATTTGAGGACAAATATATTTTAA 981
DB 191 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 132
QY 982 CACCCTCCATTAATGATATATTATTAATTTTGTTCGATCTCTTATTTTACCTCCTAAC 1041
DB 131 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 72
QY 1042 TTAATCAATTAACCAATTTTGAACCTGTTAATAATTTCTTAACCTTATTCACCTAAT 1096
DB 71 TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 17

RESULT 4
US-09-949-016-13504/c
; Sequence 13504, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13504
; LENGTH: 59519
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-13504

```
Query Match      6.3%; Score 90.2; DB 4; Length 59519;
Best Local Similarity 48.8%; Pred. No. 2.9e-09;
Matches 301; Conservative 0; Mismatches 313; Indels 3; Gaps 2;

Qy 501 ATTTTATTTTACCTCCTGCTAGATTTCGTAATACTATTTCGCAATTTATTCCTCAATTCATT 560
Db 48491 ATAATATATATATTGATGTTATATACATAATTATGTAATAATATGTTATATATATTATA 48432

Qy 561 ATTTATTTAAATTTTATATATTATTTGGATFAAAAATTCCTAATCTTACCTTTTAAAA 620
Db 48431 CATATAATGTAATAATATATATATATATATATATATATATATATATATATATATATAT 48372

Qy 621 AAGAAATTTATTTTATTTTATATATTTTAAAGATAAAAATTCCTAATCTTTCCTTTT 680
Db 48371 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48312

Qy 681 TTAATAAGAAATTTCAATTCGCTTTTCTTAAATTTAGTTTAAATTTCTATACATAATTATA 740
Db 48311 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48254

Qy 741 AAATTCGATCGGAATGATGCTGCTCAAGTCAGTCACATGAATTTTGGTGAGAAAA 800
Db 48253 TATTATGTAATAATATATATATATATATATATATATATATATATATATATATATATAT 48194

Qy 801 AATAAATAATTAACACATTTTTCGATTAATTTTATATATATATATATATATATATATATAT 860
Db 48193 ATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48134

Qy 861 TTTATTTAATGTTGCAATAATATTTTAAATTTAAATTTTTCAGCACACAATTTACACTCT 920
Db 48133 TATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48074

Qy 921 CATCAATTAATTTAATCTTATTAACATAATTAATAATTTGAGGACAAATTTTAAATC 980
Db 48073 ATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48014

Qy 981 TCACCTCCATTAATGATATATTAATTTT-TGTTCCATCTCTTATTTTCACTCTCA 1039
Db 48013 TTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 47954

Qy 1040 CATTAATCATTAACCAATTTTGAATCTGTTAAATTTCTTAATTTTCACTTATTTGTGGC 1099
Db 47953 ATATATGTAATAATATATATATAAATCCCATACGTATACATGTAACAATTTATGTC 47894

Qy 1100 TCTGGTCCATCTGGAA 1116
Db 47893 AATGTTATATATGTA 47877
```

```
RESULT 5
US-09-949-016-15129/c
; Sequence 15129, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15129
; LENGTH: 59886
; TYPE: DNA
```

```
ORGANISM: Human
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(55886)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15129

Query Match      6.2%; Score 88.8; DB 4; Length 55886;
Best Local Similarity 51.2%; Pred. No. 5.6e-09;
Matches 257; Conservative 0; Mismatches 242; Indels 3; Gaps 2;

Qy 538 TATTGCAATTTATCTCAATTTCAATTTATTTTAAATTTTATATATTTTATATATTTTATAT 597
Db 5291 TATTTTATATATATAAAAATCTTATTTTATATATATATTTTATATATATTTTATATATA 5232

Qy 598 CTAATACCTTTACTTTTTTTTAAAAAGAAATTTTAAATTTTATATATTTTATATATTTTATAGATA 657
Db 5231 ATCAATTTTATATATATAAATCTTATTTTATATATATTTTATATATTTTATATATATA 5172

Qy 658 AATTTCTAATCTTTTACTTTTTTTTTTAAAAAGAAATTTCAATTCGCTTTTTTCTTAATTTAG 717
Db 5171 ATCTTATTTTATATATATTTTATATATAAAAATCATATTTTATATATATTTTATATAT 5112

Qy 718 TTTTAAATCTATATAATATATAAAAATTCGATCGGATTTAGTGGTGTCAAGTCAAGT 777
Db 5111 TTTTATATATATGAAATCTTATTTTATATATATTTTATATATATTTTATATATATA 5054

Qy 778 CACATGAATTTTGTGGAGAAAAATTAATAATTAACACATTTTTCGATTTAATTTTATTTAT 837
Db 5053 AATCTTATTTTATATATATTTTATATATATAAATCTTATTTTATATATATTTTATAT 4994

Qy 838 ATATATATAATATAAACAATTTTATTTTAAATGTTGTCATATAATTTTAAATTAATA 897
Db 4993 ATATATAAATCTTATTTTATATATATTTTATATATATAAATCTTATTTTATATATA 4934

Qy 898 TTTGAGCACACAATTAACATCTCATCAT-TAAATTTTAACTTATACCATTAATTAATA 956
Db 4933 TTTTATATATATAAATCTTATTTTATATATATTTTATATATATAAATCTTATTT 4874

Qy 957 TGTGAGGACAAATTTTAAATCTCACCTCCATTAATGATATATTTTAAATTTTGTTC 1016
Db 4873 ATACAGCAATTTTCTAAATAAATAATACAGCCATTTTCTAAATAATATATTTTCTTT 4814

Qy 1017 GATACTTTCTTATTTTCACTCTCA 1038
Db 4813 TAAACTTTTTCCTTTTCCCA 4792
```

```
RESULT 6
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Welles, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15852
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A, T, C or G
; US-09-949-016-15852

Query Match 6.2%; Score 88.4; DB 4; Length 205044;
Best Local Similarity 49.1%; Pred. No. 9.2e-09;
Matches 288; Conservative 0; Mismatches 236; Indels 2; Gaps 2;
Qy 427 AATTATTATTTTAAAAAATGATGACGAGTGAGAGAGAGAGGTCACCGATTG 486
Db 201233 AACAAATTTTATTTGTAATCTGACTTTATCTCTGAAAGGGGATTTAGAGACAAACAC 201174
Qy 487 GTCAAGTGGACCAATTTTTTATTTTACCTCTGCTAGATTCGTAATACTATTGCAAT 546
Db 201173 TTCAATGTTTCATAGATAGGGTCAATTCCTCAATTTAAATATATTCAAATATTAATATAT 201114
Qy 547 TATCTCATTTTCAATTTATTTTAAATTTATTTTATTTTATTTTGGATAAAAAATCTAATACTT 606
Db 201113 TATATTTAAATATATAAATATATTTTATATATTTTATATTTTAAATATATATATTTTAA 201054
Qy 607 TACTTTTTTTTAAAAAGAAATTTAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 666
Db 201053 TATATTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATATTTT 200994
Qy 667 ACTTTACTTTTTTTTAAAAAGA-ATTTCATTCGGTTTTTCTTAATTTAGTTTAAAT 725
Db 200993 TATATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATAT 200934
Qy 726 CTATACATAATTATAAAAAATCTGATCGGATTTAGTGTGTCGATCAAGTCACATGAA 785
Db 200933 TTTATATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATA 200874
Qy 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATATATAA 845
Db 200873 TTTTATATATTTTATTTTAAA- TATATATATTTTATATTTTATATTTTATTTTAAATATATAT 200815
Qy 846 TAATATAAACACATTTTATTTTAAATGTCGTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 905
Db 200814 ATATTTTATATTTTATTTTAAATATATAAATATATATTTTATATATTTTATATTTTAAATTTTAA 200755
Qy 906 CAACAATTTACATCTCATTAATTTAAATTTAACTTTATTTACCATAATTTAAATTTTGGAGAC 965
Db 200754 AATATATAAATATATTTTATATTTTATATTTTAAATATATAAATATATATTTTATATATTTTATAT 200695
Qy 966 AATTATTTTAAATCTCACCTCCATTAATGCAATATTTTAAATTTT 1011
Db 200694 ATTTATATTTAAATATATAAATATATTTTATATTTTATATTTTATATTTT 200649

RESULT 9
US-09-949-016-15853/c
; Sequence 15853, Application US/09949016

Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15853
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A, T, C or G
; US-09-949-016-15853

Query Match 6.2%; Score 88.4; DB 4; Length 205044;
Best Local Similarity 49.1%; Pred. No. 9.2e-09;
Matches 288; Conservative 0; Mismatches 236; Indels 2; Gaps 2;
Qy 427 AATTATTATTTTAAAAAATGATGACGAGTGAGAGAGAGGTCACCGATTG 486
Db 201233 AACAAATTTTATTTGTAATCTGACTTTATCTCTGAAAGGGGATTTAGAGACAAACAC 201174
Qy 487 GTCAAGTGGACCAATTTTTTATTTTACCTCTGCTAGATTCGTAATACTATTGCAAT 546
Db 201173 TTCAATGTTTCATAGATAGGGTCAATTCCTCAATTTAAATATATTCAAATATTAATATAT 201114
Qy 547 TATCTCATTTTCAATTTATTTTAAATTTATTTTATTTTATTTTGGATAAAAAATCTAATACTT 606
Db 201113 TATATTTAAATATATAAATATATTTTATATATTTTATATTTTAAATATATATATTTTAA 201054
Qy 607 TACTTTTTTTTAAAAAGAAATTTAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 666
Db 201053 TATATTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATATTTT 200994
Qy 667 ACTTTACTTTTTTTTAAAAAGA-ATTTCATTCGGTTTTTCTTAATTTAGTTTAAAT 725
Db 200993 TATATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATAT 200934
Qy 726 CTATACATAATTATAAAAAATCTGATCGGATTTAGTGTGTCGATCAAGTCACATGAA 785
Db 200933 TTTATATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTATTTTAAATATATATA 200874
Qy 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATATATAA 845
Db 200873 TTTTATATATTTTATTTTAAA- TATATATATTTTATATTTTATATTTTATTTTAAATATATAT 200815
Qy 846 TAATATAAACACATTTTATTTTAAATGTCGTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 905
Db 200814 ATATTTTATATTTTATTTTAAATATATAAATATATATTTTATATATTTTATATTTTAAATTTTAA 200755
Qy 906 CAACAATTTACATCTCATTAATTTAAATTTAACTTTATTTACCATAATTTAAATTTTGGAGAC 965
Db 200754 AATATATAAATATATTTTATATTTTATATTTTAAATATATAAATATATATTTTATATATTTTATAT 200695
Qy 966 AATTATTTTAAATCTCACCTCCATTAATGCAATATTTTAAATTTT 1011
Db 200694 ATTTATATTTAAATATATAAATATATTTTATATTTTATATTTTATATTTT 200649

RESULT 10
US-09-949-016-12387/c

```
US-09-949-016-12724/c
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12387
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387

Query Match      6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;

QY 427 AATTATTATTTTAAAAAATTCATGTGACCGAGTGGTGGAGAGAGAGGCTACCGATTG 486
..DB 171660 AACAAATTTTATTTGTAATCTGACTTTATCTCCTGAAAGGGGATTTAGAGACAACAC 171601

QY 487 GTCAGTGGGCACCAATTTTTTATTTTACCTCCCTGCTAGATTCGTAAATACTATTGCAAT 546
DB 171600 TTCAATGTTTCATAGATAGGGGTCAAATTCCTCATTTAAATATATTTCAAATATTAATATAT 171541

QY 547 TAATCTCAATTCATTTTATTTTAAATTTTATTTTACCTCCCTGCTAGATTCGTAAATACTATTGCAAT 606
DB 171540 TATATTTTAAATATATAATAATATATTTTATATATTTTATATATTTTAAATATATAATATTTTA 171481

QY 607 TACTTTTTTTTAAAAAGAAATTTTAAATTTTATTTTAAATTTTATTTTATTTAGATAAAAAATTCCTAAT 666
DB 171480 TATATTTATATTTTAAATATATAATATTTTATATATTTTATATATTTTAAATATATAATATTTT 171421

QY 667 ACTTTACTTTTTTTTTTAAAAAGA-AATTCAAATTCGGTTTTTCTTAATTTAGTTTTAAT 725
DB 171420 TATATATTTTATATTTTAAAAATATATAATTTTATATATTTTATATATTTTAAAAATATATAAT 171361

QY 726 CTATACCTAATTAATAAAATTCGATCGGATTCGATCGGATTCGATCGGATTCGATCGGATTCGATCGA 785
DB 171360 TTTTATATATTTATTTTAAATATATAATATTTTATATATTTTATATATTTTAAATATATAATA 171301

QY 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATATAA 845
DB 171300 TTTTATATATTTTATATTTTAAAAA-TATATATATTTTATATATTTTATATTTTAAAAATATAT 171242

QY 846 TAATATAAACACATTTTTTATTTAATGTGTCAATAATAATTTTTTAAATTTAAATTTTCAGCA 905
DB 171241 ATATTTTATATATTTTATATTTTAAAAATATATAATAATATATTTTATATATTTTATATTTAA 171182

QY 906 CAACAATTAACCTCATCATTAATAATTTAATCTTTATACCATAATTAATAATTTGGAGGAC 965
DB 171181 AATATATATAAATATATTTTATATATTTTATATATTTTATATTTTAAATATATAATTTTATAT 171122

QY 966 AATTATTTTTTAAATCTCACCCCTCCATTTAAATGCATATTTTAAATTTT 1011
DB 171121 ATTTATATTTTAAATATATAATAATATATTTTATATATTTTATATATTTT 171076

US-09-949-016-12724/c
; Sequence 12724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12724
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

Query Match      6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;

QY 427 AATTATTATTTTAAAAAATTCATGTGACCGAGTGGTGGAGAGAGAGGCTACCGATTG 486
..DB 171660 AACAAATTTTATTTGTAATCTGACTTTATCTCCTGAAAGGGGATTTAGAGACAACAC 171601

QY 487 GTCAGTGGGCACCAATTTTTTATTTTACCTCCCTGCTAGATTCGTAAATACTATTGCAAT 546
DB 171600 TTCAATGTTTCATAGATAGGGGTCAAATTCCTCATTTAAATATATTTCAAATATTAATATAT 171541

QY 547 TAATCTCAATTCATTTTATTTTAAATTTTATTTTACCTCCCTGCTAGATTCGTAAATACTATTGCAAT 606
DB 171540 TATATTTTAAATATATAATAATATATTTTATATATTTTATATATTTTAAATATATAATATTTTA 171481

QY 607 TACTTTTTTTTAAAAAGAAATTTTAAATTTTATTTTAAATTTTATTTTATTTAGATAAAAAATTCCTAAT 666
DB 171480 TATATTTATATTTTAAATATATAATATTTTATATATTTTATATATTTTAAATATATAATATTTT 171421

QY 667 ACTTTACTTTTTTTTTTAAAAAGA-AATTCAAATTCGGTTTTTCTTAATTTAGTTTTAAT 725
DB 171420 TATATATTTTATATTTTAAAAATATATAATTTTATATATTTTATATATTTTAAAAATATATAAT 171361

QY 726 CTATACCTAATTAATAAAATTCGATCGGATTCGATCGGATTCGATCGGATTCGATCGGATTCGATCGA 785
DB 171360 TTTTATATATTTATTTTAAATATATAATATTTTATATATTTTATATATTTTAAATATATAATA 171301

QY 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATAA 845
DB 171300 TTTTATATATTTTATATTTTAAAAA-TATATATATTTTATATATTTTATATTTTAAAAATATAT 171242

QY 846 TAATATAAACACATTTTTTATTTAATGTGTCAATAATAATTTTTTAAATTTAAATTTTCAGCA 905
DB 171241 ATATTTTATATATTTTATATTTTAAAAATATATAATAATATATTTTATATATTTTATATTTAA 171182

QY 906 CAACAATTAACCTCATCATTAATAATTTAATCTTTATACCATAATTAATAATTTGGAGGAC 965
DB 171181 AATATATATAAATATATTTTATATATTTTATATATTTTATATTTTAAATATATAATTTTATAT 171122

QY 966 AATTATTTTTTAAATCTCACCCCTCCATTTAAATGCATATTTTAAATTTT 1011
DB 171121 ATTTATATTTTAAATATATAATAATATATTTTATATATTTTATATATTTT 171076

RESULT 11
```

RESULT 12

US-09-949-016-12725/c
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

Query Match 6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;
Qy 427 AATTATTATTTTAAAGAAATGATGACGAGGTGGGAGAGAGGCTACCGATTG 486
Db 171660 AACAAATTTTATTTGTAATCTGACTTATTTCTCTGAAAGGGATTTAGACAAACAC 171601
Qy 487 GTCAAGTGGCACCATTATTTTATTTTACCTCGCTAGATTCGTAATATCTATTGCAAT 546
Db 171600 TTCATGTTTCAAGTAGGCGTCAATTCCTCAATTAATATATTCGAATATATATAT 171541
Qy 547 TATCTCAATTCATTTATTTAAATTTATTTATATTTTTCGATAAAATTTCTAATCTT 606
Db 171540 TATATTTAAATATATAAATATATTTTATATATTTTATATTTTAAATATATATTTT 171481
Qy 607 TACTTTTATTTTAAAGAAATTTATTTATTTTATTTATTTTATTTTATTTTATTTTAT 666
Db 171480 TATATTTATTTTAAATATATATATTTTATATTTTATATTTTATATTTTAAATAT 171421
Qy 667 ACTTTACTTTTATTTTAAAGAA-ATTTCAATTCGGTTTCTTAAATTTAGTTTAAAT 725
Db 171420 TATATATTTATTTTAAATATATATATATTTTATATATTTTATATTTTAAATAT 171361
Qy 726 CTATACTAATTTAAATTTCTGATCGGATTTAGTGTGTCGATCAAGTCACATGAA 785
Db 171360 TTTATATATTTATTTTAAATATATATTTTATATATTTTATATTTTAAATATATA 171301
Qy 786 TTTTGTTCGAGAAAAATTTAAATTTAAACACATTTTTCGATTAATTTTATATATATAA 845
Db 171300 TTTTATATATTTTATTTTAAATTTAAATTTTATATATTTTATATTTTAAATATAT 171242
Qy 846 TAATATAACACATTTTATTTTAAATTTGTCATTAATATTTTAAATTTAAATTTTCA 905
Db 171241 ATATTTTATATTTTATTTTAAATTTAAATTTATATAAATATATTTTATATATTT 171182
Qy 906 CAACAATTTACCTCTCATCATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 965
Db 171181 AATATATATATATATTTTATATATTTTATATTTTAAATTTAAATTTATATATTT 171122
Qy 966 AATTATTTTAAATCTCACCTTCATTAATGATGATATTTTAAATTTT 1011
Db 171121 ATTTATATTTTAAATATATATAAATATATTTTATATATTTTATATTTT 171076

RESULT 13

US-09-949-016-17369
; Sequence 17369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17369
; LENGTH: 61178
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17369

Query Match 6.1%; Score 87.4; DB 4; Length 61178;
Best Local Similarity 49.9%; Pred. No. 1.1e-08;
Matches 275; Conservative 0; Mismatches 271; Indels 5; Gaps 2;
Qy 546 TTATCTCATTTTCATTTATTTTAAATTTATTTTATTTTATTTTGGATATAAAATTTCTA 605
Db 21250 TAACTCTATCTTATATAATCTTATATTTATTTATATATGCTCTTATATAATATAT 21309
Qy 606 TTACTTTTATTTTAAAGAAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTAA 665
Db 21310 ATAATCTTATCTTATATAATATAATCTTATTTATATAATAATATCTTATTTATATA 21369
Qy 666 TACTTTTACTTTTATTTTAAAGAAATTTCAATTCGCTTTTCTTAAATTTAGTTTAAAT 725
Db 21370 TCTTATCTTATATAATATAATATAATCTTATTTATATAATAATCTTATCTTATATA 21429
Qy 726 CTATACTAATTTATAAAATTTCTGATCGGATTTAGTGTGTCGATCAAGTCACATGAA 785
Db 21430 ATAATCTTATATAATAATAATCTTATTTTATATAATAATCTTATTTATATAATA 21488
Qy 786 TTTTGTTCGAGAAAAATTTAAACACATTTTTCGATTAATTTTATTTATATATAATA 845
Db 21489 ATAATCTTATCTTATATAATAATCTTATCTTATATAATAATAATCTTATCTTAT 21548
Qy 846 TAATATAACACATTTTATTTTAAATTTGTCATTAATTTTAAATTTAAATTTTCA 905
Db 21549 TTATATAATAATCTTATCTTATCTTATATAATAATAATAATAATAATAATCTTAA 21608
Qy 906 CAACAATTTACATCTCTCATTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAA 961
Db 21609 TATATAATAATAATCTTATATAATAATCTTATCTTATATAATAATAATAATAATA 21668
Qy 962 GGACAAATTTTATTTTAAATTTTCACTCCCTCCATTAATGATGATATTTTATTTTTCG 1021
Db 21669 ATATAATCTTATATTTTATATAATAATCTTATCTTATATAATAATAATAATAATA 21728
Qy 1022 TTTCTTATTTTCACTCTTAACATTAATCTTAACCAATTTTGAACCTGTTTATATTTCT 1081
Db 21729 TCTTATCTTATATAATAATGTTTATTTATATAATAATAATAATAATAATAATAATA 21788
Qy 1082 CTTATTTCACTA 1092
Db 21789 ATTAATATA 21799

RESULT 14
US-09-949-016-12776/c

Search completed: August 13, 2005, 13:15:41
Job time : 623 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 11:00:09 ; Search time 8243 Seconds
(without alignments)
1127.927 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcgcggtgttc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7305758 seqs, 3244068913 residues

Total number of hits satisfying chosen parameters: 14611516

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	137.8	9.6	4950	20	US-10-425-115-109152
C 2	136.4	9.5	4977	20	Sequence 109152, Sequence 177410, Sequence 177410, Sequence 109144, Sequence 109144, Sequence 109143, Sequence 109143, Sequence 23515, A Sequence 177422, Sequence 162914,
C 3	133.6	9.3	4623	20	US-10-425-115-109144
C 4	132	9.2	4719	20	US-10-425-115-109143
C 5	131.2	9.2	3285	19	US-10-437-963-23515
C 6	131	9.1	4064	20	US-10-425-115-177422
C 7	129.4	9.0	2460	20	US-10-425-115-162914

C 8	129.4	9.0	4530	20	US-10-425-115-109139
C 9	129.2	9.0	2373	19	US-10-437-963-78436
C 10	129.2	9.0	4641	19	US-10-437-963-78434
C 11	128.8	9.0	1155	20	US-10-425-115-177419
C 12	128.8	9.0	4884	20	US-10-425-115-109131
C 13	128.4	9.0	5433	20	US-10-425-115-177411
C 14	128	8.9	58965	17	US-10-298-122-2
C 15	126	8.8	4641	19	US-10-437-963-78438
C 16	124.8	8.7	4530	19	US-10-437-963-89508
C 17	123.5	8.6	1182	20	US-10-425-115-149480
C 18	123.2	8.6	3969	19	US-10-437-963-77068
C 19	123.2	8.6	4303	19	US-10-437-963-89510
C 20	123.2	8.6	5292	19	US-10-437-963-89510
C 21	123.2	8.6	5430	19	US-10-437-963-89526
C 22	122.2	8.5	5162	19	US-10-437-963-89553
C 23	121.6	8.5	2781	19	US-10-437-963-89549
C 24	121.4	8.5	3057	20	US-10-425-115-111975
C 25	120	8.4	4638	19	US-10-437-963-77057
C 26	120	8.4	5058	19	US-10-437-963-89540
C 27	120	8.4	5208	19	US-10-437-963-89627
C 28	120	8.4	5547	19	US-10-437-963-89646
C 29	119.8	8.4	1326	19	US-10-437-963-34349
C 30	119.8	8.4	5199	19	US-10-437-963-89684
C 31	119.6	8.3	1056	20	US-10-425-115-109133
C 32	119.4	8.3	3936	19	US-10-437-963-85379
C 33	119	8.3	5043	19	US-10-437-963-14832
C 34	119	8.3	5151	19	US-10-437-963-23475
C 35	119	8.3	9603	19	US-10-437-963-54020
C 36	119	8.3	10549	19	US-10-437-963-89620
C 37	118.8	8.3	3849	20	US-10-425-115-149479
C 38	118.4	8.3	4092	19	US-10-437-963-89558
C 39	118.4	8.3	4149	19	US-10-437-963-89589
C 40	118.4	8.3	4611	19	US-10-437-963-77761
C 41	118.4	8.3	5281	19	US-10-437-963-85626
C 42	118.4	8.3	5448	19	US-10-437-963-89685
C 43	118.4	8.3	5556	19	US-10-437-963-89725
C 44	118.4	8.3	5571	19	US-10-437-963-89616
C 45	118.4	8.3	6101	19	US-10-437-963-29850

ALIGNMENTS

RESULT 1
US-10-425-115-109152/c
; Sequence 109152, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109152
; LENGTH: 4950
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31042C.1
; US-10-425-115-109152

Query Match 9.6%; Score 137.8; DB 20; Length 4950;
Best Local Similarity 62.2%; Pred. No. 5.2e-10;
Matches 217; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
Qy 58 GATTTCATAGACTTCTATCTATCAGAGCGCTCAGAGGATCCCAATTAGTCTAA 117
Db 4492 GGTCTGCATACTCTCTGACGAGCTTGGGCTTTCTTCATATTATGAAATTTCTGAA 4433

Qy	118	AAATTATCTTTCAGTCTCGGAACCAACCACTCAGGACCCCAAAACCCGCGCGCTCAACCAACTCAG	177
Db	4432	CCTTTTCTTCAACCTCTTTCCACCATATCAGGCTTAAAGAAGTATCTTTTCTCAGGTTTCAG	4373
Qy	178	TCTAATAATAACAGAGTAGTATGACACTTATATGACCATATAGAGCCTCTGAAGGTGCCATCTAGA	237
Db	4372	ACCAATTTACGGAGTTTCGACACCGAGCTCCATATAAAGCTTCAAAGGSGTCCATCTTGA	4313
Qy	238	TGCCAGATTTGGAACCTGTTTATTGTAGGGGAACCTCACTAAACGGTAAAAAATCTCTCTCAAC	297
Db	4312	TACTTTCTTGTTAACTGTTTATTGTATGAAGAACTCCCGTAGGGGCAAAACATTCATCCCAATT	4253
Qy	298	TACCTTAGTAATAAATCACATAGCTCCAAATTCGTATCTCTTAGTATATGAATCACTTCT	357
Db	4252	TTGTTGAGAAGTCCAGACACATGCCCGCAGCATATCTTCAAGTATTTGGTTCAACCTCT	4193
Qy	358	CAAAATTGACCATCGGCTCGAGGATGGAATGCAGACCGGTGCCACCGAATT	406
Db	4192	CAGTTTGTCCACTCGTTTGGAGATGATAGGCCGAACCTGTGGAGCAGTTT	4144

```

RESULT 2
US-10-425-115-177410
; Sequence 177410, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 177410
; LENGTH: 4977
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_93387C.1
US-10-425-115-177410

```

```

RESULT 3
US-10-425-115-109144/c
; Sequence 109144, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109144
; LENGTH: 4623
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_31035C.1
US-10-425-115-109144

```

```

RESULT 4
US-10-425-115-109143/c
; Sequence 109143, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109143
; LENGTH: 4719
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

```

OTHER INFORMATION: Clone ID: MRT4577_31034C.1
US-10-425-115-109143

Query Match 9.2%; Score 132; DB 20; Length 4719;
Best Local Similarity 61.8%; Pred. No. 3.4e-09;
Matches 210; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 58 GATTTCGATGAACTCTCTATCATGAGAGCGCTGAGAGGATCCCAATTTAGTCTAA 117
DB 4261 GGTCTGATTAACCTTTTGTGAGTCTGAGCTTCTTCATATTATGATTAATTTCTGTAA 4202
QY 118 AATTATCTTCAGTCTCGAAACCAACTCAGGACCCAAACCCGCTGCTCAACCACCTCAG 177
DB 4201 CCTTTCTTCGACCTCTTTTCCACCATATCAGGCGCTTAAAGAGTATCTTTCTCCAGGTTTCA 4142

QY 178 TCTAATAATACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGA 237
DB 4141 ACCAATTTAGCGAGTTTCGACACTGTCTCCATATAAAGCTTCAAAAGGTGCCATCTTGA 4082

QY 238 TGCAGAGTTGGAACCTGTTATTTAGGCGAACTCAACTAAACGGTAAATAATCTCTCAAC 297
DB 4081 TACTTTCTGATTAAGTCTGTTATGTAATGAAATTCGCTAAGGCAACATTCATCCATT 4022

QY 298 TACCTTAGTAAATACATAGCTCCAAATCGTATCTCTTAGTATATGAATCACCTTCT 357
DB 4021 TTTGCGAGAGTCCAGAACATATGCGCGAGCATATCTCGAGTATCTGATTCACCTCT 3962

QY 358 CAAATTCACCATCGGTCTGAGGATGGAATGACGACCGGTG 397
DB 3961 CAGTCTGTCCACTGTTTGGAGTATGATAGCGCAACTGTG 3922

RESULT 5

US-10-437-963-23515/c
Sequence 23515, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 23515

LENGTH: 3285

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_28588C.1

US-10-437-963-23515

Query Match 9.2%; Score 131.2; DB 19; Length 3285;
Best Local Similarity 61.3%; Pred. No. 4e-09;
Matches 211; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 52 TTTTCAGATTGCAATGACCTTCTATCTATCAGAGAGCGCTGAGAGGATCCCAATTA 111
DB 1953 TTCTGACATCGCGTAGCTCTCTCTACTTCTGAGCGCAAGCATGTGTGCGGTATCT 1894

QY 112 GTCTAAATATCTTTCAGTCTCGGAACCACTCAGGACCCAAACCCGCTGCCTCACCA 171
DB 1893 CTGCTATCCGTTCTGTGTGTTGAACCCCAATCGGTCCTTAAATAGTCTCTCCCTTA 1834

QY 172 ACTCAGTCTAATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCA 231

DB 1833 TGGAAATCCCAACAAATGGGATTCGACATTTTCTCCATACAGTGGCTCATATGGTGCCA 1774
QY 232 TCTAGATGCGAGATTGGAACCTGTTATTTAGGCGGAACCTCAACTAAACGGTAAATAATCTCT 291
DB 1773 TTGCAATGCTAGCATGTAACATTTCTTGTAGCAAAATTTCTACGAGCGCAATGATCTCT 1714
QY 292 CTCAACTACCTTAGTAAATAAATCACAATAGCTCCAAATCGTATCTCTTAGTATATGAATCA 351
DB 1713 CCCAACTGCGCTTTCCCATGAAATTTATACATGATCTCAACATATCTCTTAGAGTTTGAATGG 1654
QY 352 CTTTCTCAATTCACCATCGTCTGAGGATGGAATGCAGACCGG 395
DB 1653 TCCGCTCAGATTGACCATCAGTCTGAGGGTGAAGGCCGTACTG 1610

RESULT 6

US-10-425-115-177422
Sequence 177422, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 177422
LENGTH: 4064
TYPE: DNA
ORGANISM: Zea mays
FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_93398C.1

US-10-425-115-177422

Query Match 9.1%; Score 131; DB 20; Length 4064;
Best Local Similarity 61.1%; Pred. No. 4.5e-09;
Matches 212; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 60 TTTGCATAAGACTTCTATCTATCAGAGAGCGCTGAGAGGATCCCAATTTAGTCTTAAA 119
DB 461 TCTGCATAACTCTTTGCGCTGCTGAGCTTCTTCATATTATGATAAATTTCTCTGAAC 520

QY 120 TTATCTTCAGTCTCGGAACCAACTCAGGACCCAAACCCGCTCACCNACTCAGTC 179
DB 521 TTTTCTTCAACCTCTTTCAACATATCAGGTCTAAAGAAAGTATCTTCTCCAGGTTCCAG 580

QY 180 TAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATG 239
DB 581 CAATTTAGCGAGTCCGACACCGCTGTCCTCATATAAGCTTCAAGGGTGCATCTTGATA 640

QY 240 CCAGATTGGAACCTGTTATTTAGGCGGAACCTCAACTAACCGTAAATAATCTCTCAACTA 299
DB 641 CTTTCTTTGGTAGCTATTATATGAAAACTCCGCTAAAGGTAAACATTCATCCATTTT 700

QY 300 CTTTAGTAATAATCAATAGCTCCAAATCGTATCTCTTAGTATATGAATCACCCTTCTCA 359
DB 701 AGTGAGAAATCCAGGACACATGCGCGCAGCATATCTTCAAGTATTTGGTCACTCTCTCA 760

QY 360 AATTGACCATCGCTGAGGATGGAATGACAGCCGCTGCCACCGATT 406

DB 761 GTTGTCCACTGTTTGGAGTATGATGCGCAACTGTTGGAGCAGTTT 807

RESULT 7

US-10-425-115-162914/c
Sequence 162914, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 162914
; LENGTH: 2460
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80152C.1
US-10-425-115-162914

Query Match          9.0%; Score 129.4; DB 20; Length 2460;
Best Local Similarity 61.4%; Pred. No. 6.6e-09;
Matches 208; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 59 ATTTGCATAGACTTCTATCTATCATCAGAGAGCGCTGCAGAGGATCCCAAAATTAGTCTAAA 118
DB 2001 ATCTGGGTAACCTCTTCTGACGAGCTTGAGCTTTCTTCAAAATTATGAATTAATTCGTTGAAC 1942

QY 119 ATTATCTTCAAGTCTCGGAAACCAACTCAGGACCCAAAACCCGCTCGCTCACCCCAACTCAGT 178
DB 1941 CTTCCTCTCTGCTCTCTTTACCATATCAGGCTGAGAAATATCTTTCACCAAGGTTTCA 1882

QY 179 CTAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGAT 238
DB 1881 CCAGTTTAACGGAGTGCAACATCGTCGTCATATAAAGCTTCGAAGGGTGCCATCTTAAT 1822

QY 239 GCAGATTGGAACCTGTATTGTAGGCGCACTCACTAAGCGTAAATAATCTCTCAACT 298
DB 1821 ACTTCTTGATAGCTATTATATATATGAGAACTCGCTAATGGTAAACAATCATCCCATTT 1762

QY 299 ACCTTAGTAATAATCATAGTCTCCAAATCGTATCTCTAGTATATGAATCACTTCTC 358
DB 1761 CTGTGGGAATCTCAGAACACATGCTCGCAACATATCTTCGAGTATCTGATTACCTCTC 1702

QY 359 AAATTGACCATCGTCTGAGGATGGAATGAGACCGGTG 397
DB 1701 AGTCTGCCACTGCTTTGAGGATGATAGGCGCAACTATG 1663

RESULT 8
US-10-425-115-109139/c
; Sequence 109139, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109139
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31030C.1
US-10-425-115-109139

Query Match          9.0%; Score 129.4; DB 20; Length 4530;
Best Local Similarity 60.8%; Pred. No. 7.7e-09;
Matches 211; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
```

```
QY 60 TTTGCATAAGACTTCTATCTATCATCAGAGAGCGCTGCAGAGGATCCCAAAATTAGTCTAAA 119
DB 4070 TCTGCATAACTCTTTTGGCGTGTGAGCTTTCTTCATATTATGGATAATTCTCTGAACC 4011

QY 120 TTATCTTCTAGTCTCGGAAACCAACTCAGGACCCAAAACCCGCTCGCTCACCAACTCAGTC 179
DB 4010 TTTTCTCTCAACCTCTTCCACCATATCAGTCTTAAGAAGATATCTTTCTCAGGTTTCAGAC 3951

QY 180 TAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATG 239
DB 3950 CAATTAGCGGAGTCCGACACCGTGTCCATATAAAGCTTCAAAGGGTGCCATCTTGATA 3891

QY 240 CCAGATTGGAACCTGTATTGTAGGCGAACTCACTAAGCGTAAATAATCTCTCACTA 299
DB 3890 CTTTCTTGATAACTATTATTATGAAAACTCCGCTAAAGGTAAACATTCGTCCTCCATTT 3831

QY 300 CCTTAGTAATAATCATAGTCTCCAAATCGTATCTCTAGTATATGAATCACTTCTCTCA 359
DB 3830 AGTGAAATCCAGGACACATGCCCGTAGCATATCTTCAAGTATTTGGTTCACTCTCTCA 3771

QY 360 AATTGACCATCGCTCTGAGGATGGAATGAGACCGGTGCCACCGATT 406
DB 3770 GTTTGCTCACTGTTTTGAGGATGATAGGTGCAACTGTGGAGCAGTTT 3724

RESULT 9
US-10-437-963-78436/c
; Sequence 78436, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78436
; LENGTH: 2373
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2373)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78243C.1
US-10-437-963-78436

Query Match          9.0%; Score 129.2; DB 19; Length 2373;
Best Local Similarity 60.1%; Pred. No. 7e-09;
Matches 215; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 38 CTTGATATCTATGATTTTCAGATTTTCATTAAGACTTCTATCTATCAGAGACGCTCAG 97
DB 2166 CTCCTAATTCCTGCTTTTGACATCGGCGTAGCTCTTCTGCTACTTTTGAGCTGGAAGCAT 2107

QY 98 AGGATCCCAAAATTAGTCTAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
DB 2106 GTGTTGTGTTTCTGCTATCCGTTCTGTGGTCTGTGTACCAATCGGTCCTCTAAAT 2047

QY 158 CCGTCGCTCACCAACTCAGTCTAAATATAACAGAGTATGACATTTATGACATATAGAGC 217
DB 2046 AGCTCTCTCCCTATGGAATCCCAACACATATAGGATCGACACTTTCTTCCATAAAGTGC 1987

QY 218 CTGTAAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATTGTAGGCGAACTCAACTAA 277
```

Db 1986 CTCGTGTTGTCACCTTGAATGCTAGCATGGAACACTATTGTTGAAGTAAATCTTCAAG 1927
Qy 278 CGGTAAATAATCCTCTCACTACCTTAGTAAATACATAGCTCCAAATCGTATCCTC 337
Db 1926 CGGCAATATGATCTCCCAACTGCTTTCCATGAAGGATACATGATCTCAACATATCCTC 1867
Qy 338 TAGTATATGAATCACTCTTCAAAATGACCATCGGTCTGAGGATGGAATGAGACCGG 395
Db 1866 TAAAGTTTGAATAGTCGCTCAGATTGACCATCAGCTGAGGGTGAAGAGCGGTACTG 1809

RESULT 10

US-10-437-963-78434/c
; Sequence 78434, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78434
; LENGTH: 4641
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78241C.1
US-10-437-963-78434

Query Match 9.0%; Score 129.2; DB 19; Length 4641;
Best Local Similarity 60.1%; Pred. No. 8.2e-09;
Matches 215; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 38 CTTGATATCTATGATTTTCAGATTGCAATAGACTTCTATCTATCAGAGAGCGCTGCAG 97
Db 4032 CTCCTAATCTCGTCTTGACATCGCGTAGCTCTTCGCTACTTTCGAGCTGAAGAT 3973
Qy 98 AGGATCCCAATATTAGTCTAAATATTCTTCAGTCTCGGAACCAACTCAGGACCCAAAC 157
Db 3972 GTGTTGTCGATTCTACTATCCGTTCTGTGTCGTTGTACCCCAATCGGTCTCTAAAT 3913
Qy 158 CCGTGGCTCACCCTCACTCAGTCTAATATAACAGAGATGACACTTATGACCATATAGGC 217
Db 3912 AGTTCTCTCCCTATGGAATCCCAACACAAATGGGATCGACACTTTCTTCATAAGTGC 3853
Qy 218 CTCGTAGTCCCTAGATGCGCAGATTGGAACGTTATTGTAGGCGCACTCAACTAA 277
Db 3852 CTCATATGGTGCCATTTGAATGCTAGCATGGAACACTATTGTTGAAGCAATCTTACAG 3793
Qy 278 CGGTAAATAATCCTCTCACTACCTTAGTAAATACATAGCTCCAAATCGTATCCTC 337
Db 3792 CGGCAATGATCCTCCCACTGCTTCCATGAAGGATACATGATCTCAACATATCCTC 3733
Qy 338 TAGTATATGAATCACTCTTCAAAATGACCATCGGTCTGAGGATGGAATGAGACCGG 395
Db 3732 TAAAGTTTGAATAGTCCGCTCAGATTGACCATCAGTCTGAGGTGAAGAGCGGTACTG 3675

RESULT 11

US-10-425-115-177419/c
; Sequence 177419, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 177419
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_93395C.1
US-10-425-115-177419

Query Match 9.0%; Score 128.8; DB 20; Length 1155;
Best Local Similarity 61.2%; Pred. No. 6.7e-09;
Matches 208; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 58 GATTTCATAGACTTCTATCTATCAGAAGAGCGCTCAGAGGATCCCAAATTAAGTCTAA 117
Db 697 GGTCTGATAACTCTTTTGGCGAGCTCGAGCTTCTTCATATTAATTAATTTCTGAA 638
Qy 118 AATATCTCTCAGTCTCGGAACCAACTCAGGACCCAAACCCGTCGCTCACCCAACTCAG 177
Db 637 CTTTCTTCGACCTCTTTTACCATAATCAGGCTTAAGAGTACCTTTCTCCAGTTTCAG 578
Qy 178 TCTAATAATAACAGAGTATGACATTATGACCATATAGAGCTCTGTAAGTGCCATCTAGA 237
Db 577 ACCAATTTAGCGAGTTTCGACACCGTTGCCCATACAAAGCTTCAAAAGGTGCCATCTTGA 518
Qy 238 TGCAGATTGGAACCTGTTATTGTAGGCGAAGCTCAACTAAGCGTAAATAATCTCTCAAC 297
Db 517 TACTTTCTTGATAACTGTTATTGTATGAAAACCTCCGTAAGGGCAACATTCATCCCAT 458
Qy 298 TACTTATGATAAATCACAATAGCTCCAAATCGTATCTCTAGTATATGAATCACCTTCT 357
Db 457 TCTGTGAGAGTCCGAAACACATCGCCGCGAGCATATCTTCAGTATTGATTGACCTCT 398
Qy 358 CAAATTGACCATCGGTCTGAGGATGGAATGAGACCGGTG 397
Db 397 CAGTCTGCCACTGTTTGGAGGATGATAAGCCGAAGTGTG 358

RESULT 12

US-10-425-115-109131/c
; Sequence 109131, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109131
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31023C.1
US-10-425-115-109131

Query Match 9.0%; Score 128.8; DB 20; Length 4884;
Best Local Similarity 61.2%; Pred. No. 9.5e-09;
Matches 208; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 78438
 ; LENGTH: 4641
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_78245C.1
 US-10-437-963-78438

Query Match 8.8%; Score 126; DB 19; Length 4641;
 Best Local Similarity 59.5%; Pred. No. 2.3e-08;
 Matches 213; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy	38	CTTCATATCATGATTTTCAGATTTCATAGACTTCTATCTATCAGAGACGCTGCAG	97
Db	4161	CTCTAATTCCTGCTTCTGACATCGGCGTAGCTCTTCTGCTACTTTGAGCTGCAAGCAT	4102
Qy	98	AGGATCCCAAAATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAAAC	157
Db	4101	GTGTTGCTGATTTCTGCTATCCGTTCTGTGCTCTGTTGTACCCACCGGGTCTTAAAT	4042
Qy	158	CCGTGCTCACCCAACTCAGTCTTAATATAACAGAGATGACACTTATGACCATATAGAC	217
Db	4041	AGCTCTCTCCCTATGGAATCCCAACACATGGGGATCGACACTTCTTCCATAAAGTGC	3982
Qy	218	CTCGTAAGTGCCATCTAGATGCCAGATTGGAACCTGTTATTTAGGGGACTCAACTAA	277
Db	3981	CTCATATGGTGCCATTTGAATGCTAGCATGGAAACTATTGTTGAAGAAAATTTCAAG	3922
Qy	278	CGGTAAAAAATCCCTCTCAACTACCTTAGTATAAATCACATAGCTCCAAATCGTATCTC	337
Db	3921	CGGCAATGATCCCTCCCACTGCTTTCCATGAAAGGATACATGATCTCAACATATCTC	3862
Qy	338	TAGTATATGAATCACCTTTCTCAAAATGACCATCGGTCTGAGGATGGAATGCAGACCGG	395
Db	3861	TAAAGTTTGAATAGTCCGCTCAGATTGACCATCAGTCTGAGGGTGAAGGCGGTACTG	3804

Search completed: August 13, 2005, 22:17:26
 Job time : 8245 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 12, 2005, 18:33:08 ; Search time 1232 Seconds
(without alignments)

6885.543 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcagcgtgtgc.....tgagagaaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1980s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1433	100.0	1433	6	ABK15699 Cotton fi
2	120	8.4	2000	8	ADA71742 Rice gene
3	113.8	7.9	31491	4	ADL10203 Pepper Bs
4	113.8	7.9	31491	4	AAF63301 Pepper Bs
5	113.8	7.9	31491	6	AAK98863 Nucleic a
6	107.4	7.5	8056	8	ABZ10246 Haematopo
7	105.8	7.4	8056	8	ABZ10100 Haematopo
8	105	7.3	8222	8	ACF62816 Colon can
9	105	7.3	11222	10	ADB54318 Pretreat
10	105	7.3	59590	3	AAF22281 BAC conta
11	104.8	7.3	158001	12	ADL17884 Human pho
12	104.4	7.3	4985	6	ABQ75107 Anopheles
13	104.4	7.3	4985	10	ACF79720 Mosquito
14	103.8	7.2	163319	3	AAF22306 Arabidops
15	103.4	7.2	8222	8	ACF62794 Colon can
16	103.4	7.2	11222	10	ADB54190 Pretreat
17	102.4	7.1	2000	12	ADJ41131 Plant cdn
18	102.2	7.1	109973	3	AAF22298 BAC conta
19	101.6	7.1	8056	8	ABZ10246 Haematopo
20	101	7.0	110000	3	AAF22305_08 Continuation (9 of

C 21	100	7.0	8056	8	ABZ10100 Haematopo
C 22	99.2	6.9	79122	3	AAF22294 BAC conta
C 23	99	6.9	101786	3	AAF22293 BAC conta
C 24	97.6	6.8	86584	3	AAF22292 BAC conta
C 25	97.4	6.8	110000	3	AAF22302_2 Continuation (3 of
C 26	96.2	6.7	15548	6	ABU34155 Human imm
C 27	96.2	6.7	92584	3	AAF22288 BAC conta
C 28	96	6.7	18154	6	ABU32254 Human imm
C 29	95.4	6.7	6109	6	ABU32326 Human imm
C 30	95.4	6.7	6109	6	AAE61077 Human gen
C 31	94.6	6.6	6963	6	ABU32979 Human imm
C 32	94.4	6.6	419	8	ABX46069 Bovine ES
C 33	94.4	6.6	5286	13	ADS89278 Oligonuc
C 34	94.4	6.6	5286	13	ADS89552 Oligonuc
C 35	93.6	6.5	9642	6	ABU32357 Human che
C 36	91.8	6.4	6216	6	ABK39932 Human che
C 37	91.8	6.4	6216	6	ABU70139 Chemical
C 38	91.8	6.4	7167	6	ABU32400 Human imm
C 39	91.2	6.4	6048	6	ABQ67002 Human ang
C 40	91.2	6.4	12237	6	ABU34358 Human imm
C 41	90.8	6.3	7814	4	AAE46530 Tumour su
C 42	90.6	6.3	6419	6	ABU32267 Human imm
C 43	90.2	6.3	778	6	ABQ15588 Oligonuc
C 44	90.2	6.3	778	6	ABQ15589 Oligonuc
C 45	90.2	6.3	6175	6	ABU33307 Human imm

ALIGNMENTS

RESULT 1

ABK15699

ID ABK15699 standard; DNA; 1433 BP.

XX

AC ABK15699;

XX

DT 21-MAY-2002 (first entry)

XX

DE Cotton fibre-specific beta tubulin, CFTUB2, promoter fragment.

XX

KW Cotton; ds; CFTUB2; fibre-specific beta tubulin; transgenic. promoter;
KW plant; anthocyanin gene; silk protein gene; cotton fibre strength;
KW polyhydroxybutyrate.

XX

OS Gossypium hirsutum.

XX

PH Key Location/Qualifiers

FT promoter

FT 449..1433

FT /*tag= a

FT /note= "Cotton fibre-specific promoter. This sequence is

FT specifically claimed in claim 3"

XX

PN WO200210377-A1.

XX

XX

PD 07-FEB-2002.

XX

XX

PF 01-AUG-2000; 2000WO-SG000111.

XX

PR 01-AUG-2000; 2000WO-SG000111.

XX

PA (MOLE-) INST MOLECULAR AGROBIOLOGY.

XX

XX

PI Cai L, Li X, Cheng N, Liu J;

XX

XX

DR WPI; 2002-217120/27.

XX

PT New fiber-specific beta-tubulin promoter from cotton for controlling gene

PT expression in cotton fibers and creating transgenic plants, in particular

PT cotton plants, having altered fiber characteristics.

XX

XX Claim 2; Fig 2; 30pp; English.

XX

CC The invention relates to a promoter that is cotton fibre-specific,

comprising the promoter of the cotton beta-tubulin gene CFTUB2. The promoter is fibre-specific in cotton and controls specific gene expression at the transcriptional level in cotton fibres and is useful for improving cotton fibres to create new cotton varieties with high fibre quality and yield by gene manipulation. The promoter is useful for creating transgenic plants, in particular cotton having altered fibre characteristics, and permits selective expression of a transgene in the cotton fibre, permitting greater latitude in the types of transgenes employed. Examples of expression of desirable genes in cotton fibre, but not in other parts of the cotton plants, include anthocyanin genes for coloured cotton, silk protein genes from silk worm or spiders for increased strength of cotton fibre, and biosynthesis of polyhydroxybutyrate in cotton fibre for improved thermal properties and insulating characteristics. The promoter can improve cotton fibres to create new cotton varieties with higher fibre quality and yield. The present sequence is the CFTUB2 promoter fragment

XX Sequence 1433 BP; 466 A; 285 C; 180 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 1433; DB 6; Length 1433;
Best Local Similarity 100.0%; Pred. No. 8.1e-189; Mismatches 0; Gaps 0;
Matches 1433; Conservative 0; Indels 0;

QY 1 ACTATAGGCGACGCGTGTGACGCGCCCGGGCTGGTCTTGATATCATGATTTTCAGAT 60
DB 1 ACTATAGGCGACGCGTGTGACGCGCCCGGGCTGGTCTTGATATCATGATTTTCAGAT 60
QY 61 TTGCATAAGACTTCTATCTATCAGAGAGCGCTGCGAGAGGATCCCAATTAGTCTAAAT 120
DB 61 TTGCATAAGACTTCTATCTATCAGAGAGCGCTGCGAGAGGATCCCAATTAGTCTAAAT 120
QY 121 TATCTTCAGTCTCGGAAACCACTCAGGACCCCAAAACCGGTGCTCACCACCACTCAGTCT 180
DB 121 TATCTTCAGTCTCGGAAACCACTCAGGACCCCAAAACCGGTGCTCACCACCACTCAGTCT 180
QY 181 AATATAACAGAGATGACACTTATGACCATATAGAGCGCTCGTAGGTGCCATCTAGATGC 240
DB 181 AATATAACAGAGATGACACTTATGACCATATAGAGCGCTCGTAGGTGCCATCTAGATGC 240
QY 241 CAGATTGGAACTGTTATTGTAGCGCACTCACTAAGCGTAAAGAAATCCTCTCAACTAC 300
DB 241 CAGATTGGAACTGTTATTGTAGCGCACTCACTAAGCGTAAAGAAATCCTCTCAACTAC 300
QY 301 CTTAGTAAATAAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCCTCTCAA 360
DB 301 CTTAGTAAATAAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCCTCTCAA 360
QY 361 ATTGACCATCGGTCTGAGGATGGAATGACAGCCGGTGCCACCGATTCTAATGGTACCT 420
DB 361 ATTGACCATCGGTCTGAGGATGGAATGACAGCCGGTGCCACCGATTCTAATGGTACCT 420
QY 421 ATAAAAAATTTATTTTTTAAAAAATGATGTGACAGTGGTGGAGAGAGGCTAC 480
DB 421 ATAAAAAATTTATTTTTTAAAAAATTTGATGTGACAGTGGTGGAGAGAGGCTAC 480
QY 481 CGATTGCTCAAGTGGCCCAATTTTTTATTTTACCTCCTCGCTAGATTGCTAAATACTAT 540
DB 481 CGATTGCTCAAGTGGCCCAATTTTTTATTTTACCTCCTCGCTAGATTGCTAAATACTAT 540
QY 541 TGCATTATCTCAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
DB 541 TGCATTATCTCAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
QY 601 ATACTTTACTTTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
DB 601 ATACTTTACTTTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
QY 661 TCTAATACTTTACTTTTTTAAAAAGAAATTTCAATTGCGTCTTTCTTTAATTTAGTTTT 720
DB 661 TCTAATACTTTACTTTTTTAAAAAGAAATTTCAATTGCGTCTTTCTTTAATTTAGTTTT 720
QY 721 TAAATCTTATCTAATTAATAAAATTTCTGATCGGATTAGTGTGGTCAAAAGTCAAGTCAAC 780
DB 721 TAAATCTTATCTAATTAATAAAATTTCTGATCGGATTAGTGTGGTCAAAAGTCAAGTCAAC 780

DB 721 TAAATCTTATCTAATTAATAAAATTTCTGATCGGATTAGTGTGGTCAAAAGTCAAGTCAAC 780
QY 781 ATGAATTTTGTGGAGAAAAATAAAAAATTAACACATTTTTTCGATTAAATTTATTATATA 840
DB 781 ATGAATTTTGTGGAGAAAAATAAAAAATTAACACATTTTTTCGATTAAATTTATTATATA 840
QY 841 TATAATAATATAAACACATTTTTTATTTAATGTGTCAATAATTTTTTAAATTTAAATTTT 900
DB 841 TATAATAATATAAACACATTTTTTATTTAATGTGTCAATAATTTTTTAAATTTAAATTTT 900
QY 901 CAGCACAAATATACACTCTCATTAATTAATTTTATTTACCATTAATTTAAATTTTGG 960
DB 901 CAGCACAAATATACACTCTCATTAATTAATTTTATTTACCATTAATTTAAATTTTGG 960
QY 961 AGGACAATTTATTTTAAATCTCACCTCCATTAATGTCAATATTTATTTTGTTCGATA 1020
DB 961 AGGACAATTTATTTTAAATCTCACCTCCATTAATGTCAATATTTATTTTGTTCGATA 1020
QY 1021 CTTCTTATTTCACTCCTAACATTAATTAATTAATTTTGAATTTTGAATTTTCTTA 1080
DB 1021 CTTCTTATTTCACTCCTAACATTAATTAATTAATTTTGAATTTTGAATTTTCTTA 1080
QY 1081 ACTTATTCATTTGTGGCTCTGGTCCATCTGGAAGGCCACCGTCCAGGCTGTCCAAC 1140
DB 1081 ACTTATTCATTTGTGGCTCTGGTCCATCTGGAAGGCCACCGTCCAGGCTGTCCAAC 1140
QY 1141 CACACTTTGCCAGCTCATCAATTTCCAGTAACTACATTTGTTCAGTTACTAAGCAATCCC 1200
DB 1141 CACACTTTGCCAGCTCATCAATTTCCAGTAACTACATTTGTTCAGTTACTAAGCAATCCC 1200
QY 1201 AATTTCAAAAATTTCAATTTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTAAACC 1260
DB 1201 AATTTCAAAAATTTCAATTTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTAAACC 1260
QY 1261 CAGCTCAACCTGCGTCAATTTAACCGGAAATCTTTTAACTCCTCTATATAACCCAAACCA 1320
DB 1261 CAGCTCAACCTGCGTCAATTTAACCGGAAATCTTTTAACTCCTCTATATAACCCAAACCA 1320
QY 1321 CTTCTCATCACTTTCCCATAAAAAGAAATTTCCGGAATTTCTTATTTCTTTTATATTTT 1380
DB 1321 CTTCTCATCACTTTCCCATAAAAAGAAATTTCCGGAATTTCTTATTTCTTTTATATTTT 1380
QY 1381 CTTCTCAATTTCCCGTCACTTTCCGAGAAAAATGAGAGAAATCTTTCAATC 1433
DB 1381 CTTCTCAATTTCCCGTCACTTTCCGAGAAAAATGAGAGAAATCTTTCAATC 1433

RESULT 2

ADA71742
ID ADA71742 standard; DNA; 2000 BP.

XX ADA71742;

XX AC

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5067.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI

```

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPT; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5067; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 569 A; 433 C; 409 G; 589 T; 0 U; 0 Other;
Query Match 8.4%; Score 120; DB 8; Length 2000;
Best Local Similarity 60.4%; Pred. No. 4.9e-08;
Matches 198; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 63 GCATAAGACTTCTATCTATCAGAGACGCGCTGAGAGGATCCCAAATAGTCTAAATTA 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 GCGTAATCTTCTGACGGTTCTGTGTCTCTTCAACGCTTCTCGTATCAGCTTAACCTGT 480
-QY 123 TCCTCAGTCTCGGAACCACTCAGGACCCAAACCGTCGCTCACCACCACTCAGTCTAA 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 TCCTCTCTGCTGACTTGATGTATGTCAGGACCGGAATATCAGCGCTTCGCCCACTCATTCAG 540
QY 183 TATAACAGAGTATGACACTTATGATACATATAGAGCCTCGTAAGGTGCATCTAGATGCCA 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 CATAGGGAGTGGCGCACTTCTTCGACACATTCGCTTGTTCGGGACATCTGAATGCTG 600
QY 243 GATTGGAACCTGTTATTGTAGGCGAATCACTAACCGTAAATAATCCTCTCAACTACCT 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 GCCTGATAGCTGTTGTTGTATGAGAACTCAGCATACGCGCAACCAACGATCCCAAGTCCCT 660
QY 303 TAGTAATAATACATAGCTCCAAATCGTATCCTCTAGTATATGAATCACTTCTCAAT 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
661 TCGAAATCCAGGCGACACGACGATGATGCTTCTTAGGATTTGGTTTACCCCTTCTGTC 720
QY 363 TGACCATCGGTCTGAGGATGGAATGCCAG 390
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
721 TGACCATCGGTCTGCGGATGGTAGGCCG 748
RESULT 3
AAD10203
ID AAD10203 standard; DNA; 31491 BP.
XX
AC AAD10203;
XX
DT 24-SEP-2001 (first entry)
XX
DE Pepper Bs2 gene.
XX
KW Bs2 gene; pepper; resistance gene; plant pathogen; virulence gene;
KW bacterial spot disease; Xanthomonas campestris pv. vesicatoria; AvrBs2;
KW hypersensitive response; transgenic plant; tomato; tobacco; rice; corn;
KW wheat; ds.
XX
OS Capsicum annuum.
XX
FH Key Location/Qualifiers
FT promoter 1..502
FT /*tag= a

```

```

FT exon 503..554
FT /*tag= b
FT /number= 1
FT /note= "This region contains a portion of 5' untranslated
FT region (5' UTR)"
FT intron 555..1439
FT /*tag= c
FT /number= 1
FT /note= "The 5' untranslated region (5' UTR) continues in
FT this region"
FT exon 1440..4162
FT /*tag= d
FT /number= 2
FT /note= "The region 1440-1479 contains 5' UTR which is
FT followed by the coding region"
FT CDS 1480..31219
FT /*tag= e
FT /product= "Bs2 protein"
FT intron 4163..31184
FT /*tag= f
FT exon 31185..31216
FT /*tag= g
FT /number= 3
FT 3'UTR 31220..31491
FT /*tag= h
FT US6262343-B1.
FT 17-JUL-2001.
FT 23-JUL-1999; 99US-00360186.
FT 23-JUL-1998; 98US-0093957P.
FT (REGC ) UNIV CALIFORNIA.
FT Staskawicz BJ, Dahlbeck D, Tai TH;
FT WPI; 2001-450496/48.
FT P-PSDB; AAE05409.
FT Nucleic acid molecules encoding Bs2 protein, useful for producing
FT transgenic plants having resistance to the plant pathogen Xanthomonas
FT campestris.
FT Claim 7; Col 21-50; 37pp; English.
XX
CC The present sequence is Bs2 gene from pepper. The Bs2 gene is shown to
CC confer resistance to plant pathogen Xanthomonas campestris pv.
CC vesicatoria which causes bacterial spot disease. The Bs2 protein has Bs2
CC biological activity, i.e., when co-expressed in a plant with a X.
CC campestris AvrBs2 gene product, it produces a localised hypersensitive
CC response. The protein includes a nucleotide binding motif and leucine
CC rich repeats of the type found in other plant resistance genes. The Bs2
CC nucleic acid molecule is useful for producing transgenic plants such as
CC pepper, tomato, tobacco, broccoli, cauliflower, cabbage, cowpea, grape,
CC canola, bean, soybean, rice, corn, wheat, barley, citrus, cotton, cassava
CC and walnut, that are resistant to plant pathogen Xanthomonas campestris.
CC The transgenic plants produced using Bs2 molecule develop a
CC hypersensitive response to the pathogen at the site of inoculation and
CC show an enhanced resistance to systemic infection. The Bs2 nucleic acid
CC molecule is also useful as polymerase chain reaction (PCR) primers for
CC amplifying portions of Bs2 nucleic acid molecule, as sequencing primers
CC to verify the authenticity of an amplified molecule, and as hybridisation
CC probes
XX
SQ Sequence 31491 BP; 9652 A; 6344 C; 5992 G; 9503 T; 0 U; 0 Other;
Query Match 7.9%; Score 113.8; DB 4; Length 31491;
Best Local Similarity 59.5%; Pred. No. 2.7e-07;
Matches 210; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

```



```
QY 560 TATTATTAAATTAATTAATTAATTTGGATAAAATCTTAATCTTATCTTTTAA 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TATATATTTAAATTTTATATTTTATTTATTTATTTATTTATTTATTTT-- 172

QY 620 AAGAATTTATTAATTTATTAATTAATTTAGATAAAATCTTAATCTTTT 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 TATTATTTTAAATTTAAATTTTAAATTTATTTATTTATTTATTTATTTAA 232

QY 680 TTTAAAGAAATTCATTCGGTTTTCTTAATTTAGTTTAAATCTATCTAATATA 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 ATTTTAAAAAATTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAAATTTT 292

QY 740 AAAATTCGATCGGATTAGTGTGTCGCAAGTCAAGTCAATGAATTTGTGGAGAA 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ATTTTAAAAATTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTAATA 352

QY 800 AATAAAAAATTAACACATTTTTCGATTAATTTAGTTTAAATCTATCTAATATA 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 ATTTATTAATTTATTTATTTTAAAAATTAATAAAAAATTAATTTAAAAAATA 472

QY 916 ACTCTCATCTAAATTTTAAATCTTATTAACATAATTTAAAAATTTGTGGAGCAATTTT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 AATATATTTAAATTAATAAAAAATTTAAAAATTTTATAATTTTAAAAATTTATATAT 532

QY 976 TAATCTCACCCTCCATTAATGATATATTAATTTTGTTCGATCTCTTATCTCACTC 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 TAAATTTAAATTTTATATAATTTATTAATAATGTTTTTAAATTAATTTATTTAATA 592

QY 1036 CTAACTTAATCATTAACCAATTTTGAACGTGTAATAATTTCTTAATCTTACTTA 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 ATAATATTTAATAAAAAATTAATTAATAATTTTAAAAATTTTAAAAAATTTTAA 649

RESULT 7
ABZ10100
ID ABZ10100 standard; DNA; 8056 BP.
AC ABZ10100;
XX
DT 16-JAN-2003 (first entry)
DE
XX
KW Haematopoietic cell proliferation disorder related DNA sequence #240.
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwöbe I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
```

```
XX Claim 28; SEQ ID NO 240; 117pp; English.
XX
XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 7.4%; Score 105.8; DB 8; Length 8056;
Best Local Similarity 50.9%; Pred. No. 3.9e-06;
Matches 304; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY 500 AATTTTATTTTACCTCGCTAGATTGTAATACTATTGCAATTTATCTCATTTTCAT 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 AATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 114

QY 560 TATTATTTAAATTTATTTATTTATTTGGATAAAAAATCTTAATCTTACTTTTAA 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TATTATTTTAAATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 172

QY 620 AAGAATTTATTTATTTATTTATTTATTTAGATAAAAAATCTTAATCTTACTTTT 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 TATTATTTTAAATTTTAAATTTTAAATTTTATTTATTTTATTTTATTTTATTTTAA 232

QY 680 TTTAAAGAAATTTCAATTCGTTTTCTTAATTTAGTTTAAATCTATCTAATATA 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 ATTTTAAAAATTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAAATTTTATTT 292

QY 740 AAAATTCGATCGGATTAGTGTGTCGCAAGTCAAGTCAATGAATTTGTGGAGAA 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 ATTTTAAAAATTTATTTTAAAAAATAATTAATAAAAAATTAATAAAAAATTAATA 352

QY 800 AATAAAAAATTAACACATTTTTCGATTAATTTAGTTTAAATCTATCTAATATA 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 ATTTAAAAAATTAATAAAAAATTAATAATTTTATTAATTTTAAAAATTTTAAATAT 412

QY 860 TTTTATTTAATGT---TGTCAATAATATTTTAAATTTAAAAATTTTCAAGCAACAATTC 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 ATTTATTAATTTATTTATTTTAAAAATTAATAAAAAATTAATTTAAAAAATA 472

QY 916 ACTCTCATCTAAATTTTAAATCTTATTAACATAATTTAAAAATTTGTGGAGCAATTTT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 AATATATTTAAATTAATAAAAAATTTAAAAATTTTATAATTTTAAAAATTTATATAT 532

QY 976 TAATCTCACCCTCCATTAATGATATATTAATTTTGTTCGATCTCTTATCTCACTC 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 TAAATTTAAATTTTATATAATTTATTAATAATGTTTTTAAATTAATTTATTTAATA 592

QY 1036 CTAACTTAATCATTAACCAATTTTGAACGTGTAATAATTTCTTAATCTTACTTA 1092
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 ATAATATTTAATAAAAAATTAATTAATAATTTTAAAAATTTTAAAAAATTTTAA 649
```

[illegible]

XX The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents, distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytosinatic activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The pNA (peptide nucleic acid)-oligomers are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the pretreated genomic DNA
CC region of the invention. This sequence is not shown within the
CC specification but is taken from Wipoweb.
XX

SQ Sequence 11222 BP; 2767 A; 0 C; 2498 G; 5957 T; 0 U; 0 Other;

Query Match 7.3%; Score 105; DB 10; Length 11222;
Best Local Similarity 49.9%; Pred. No. 4.8e-06;
Matches 345; Conservative 0; Mismatches 340; Indels 6; Gaps 3;
QY 412 ATGGTACCTATAAAAAATTAATTTTAAAAAAATTTGATGACCAAGTGGTGGAGAGA 471
DB 7223 ATGGTGTGTTGCTGAAGTTGAATGAATAAATGTTTATGAGTAATAGGGTGGATGGT 7282
QY 472 GAGGCTACCGATGGTCAAGTGGCCCAATTTTTTATTTTACCTCGCTAGATTCGT 531
DB 7283 TTTGTAAGAAAGAGAGATGGAGGATGATTAGTTTTTATTTTATTTATTTATTTT 7342
QY 532 AATATCATTTGCTATTTATCTCATTTCAATTTATTTTATTTTATTTTATTTTATTTT 589
DB 7343 AATTTATTTAGGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7402
QY 590 TAAAAATTTCAATACCTTTACTTTTTTTTAAAAAGAAATTTATTTAAATTTATTTATTTATTT 649
DB 7403 TATTTATTTATTTGTTATTTTAAATTTATTTATTTATTTATTTTGTTTATTTATTTATTTA 7462
QY 650 TAGATAAAAATCTAATACCTTTACTTTTTTTTAAAAAGAAATTTCAATTCGTTTCTTCT 709
DB 7463 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7522
QY 710 TAAATTTAGTTTAAATCTATCTAATTTATATAAAATTTCTGATCGATTAGTGTGCTCA 769
DB 7523 TGAATTTTAAATTTATTTATTTATTTATTTATTTATTTTGTTTATTTAGTATTTA 7582
QY 770 AGTCAAGTCAATGAATTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTA 829
DB 7583 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7642
QY 830 TTTTATTAATATAAATAAATAAACAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 889
DB 7643 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7702
QY 890 AATTTAA--ATTTTACGACACAATTTACACTCTCATCATTAATAATTTAACTTTATTTACCAT 947
DB 7703 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7762
QY 948 AATTTAAATTTGAGGACAATTTATTTTAAATCTCACCCCTCCATTAATGATATTTATTA 1007
DB 7763 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7822
QY 1008 TTTTGTTCGA--TACTTCTTATTTCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1065
DB 7823 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7882
QY 1066 TGTATTAATTTCTTAATCTTATTTACTATTGT 1096
DB 7883 TATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7913

AAF22281/c

ID AAF22281 standard; DNA; 59590 BP.

XX AAF22281;

AC AAF22281;

XX 20-MAR-2001 (first entry)

XX BAC containing repeats from centromeres 1-4 #4.

XX Centromere; microsome; vector; ds.

XX Arabidopsis thaliana.

XX WO200055325-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-US007392.

XX 18-MAR-1999; 99US-0125219P.

XX 01-APR-1999; 99US-0127409P.

XX 18-MAY-1999; 99US-0134770P.

XX 13-SEP-1999; 99US-0153584P.

XX 17-SEP-1999; 99US-0154603P.

XX 16-DEC-1999; 99US-0172493P.

XX (UYCH-) UNIV CHICAGO.

XX Preuss D, Copenhaver G, Keith K;

XX WPI; 2000-587529/55.

XX Recombinant DNA construct comprising a plant centromere, useful for

XX producing stably inherited microsome which can serve as vectors for the

XX construction of transgenic plant and animal cells expressing selected

XX proteins such as hormones, enzymes, interleukins, clotting factors,

XX cytokines, antibodies, and growth factors

XX Sequence 59590 BP; 17614 A; 12031 C; 13575 G; 16118 T; 0 U; 252 Other;

XX Claim 102; Page 351-364; 1449pp; English.

XX The present invention relates to a recombinant DNA construct of a plant

XX (Arabidopsis thaliana) centromere. The constructs are useful for

XX producing stably inherited microsome which can serve as vectors for the

XX construction of transgenic plant and animal cells expressing selected

XX proteins such as hormones, enzymes, interleukins, clotting factors,

XX cytokines, antibodies, and growth factors

XX Query Match 7.3%; Score 105; DB 3; Length 59590;

XX Best Local Similarity 57.4%; Pred. No. 4.1e-06;

XX Matches 189; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 59 ATTTCATTAAGACTTCTTATCTATCAGAAGCGCTCGAGAGGATCCAAATTAAGTCTAAA 118

DB 20299 ATCCGTAATTAATCTTCTGTCGGTTCTGAGATACCTTTAACTTAACCTTAAGAAAGTTCAT 20240

QY 119 ATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAACCCGTCGCTCACCACCACTCAGT 178

DB 20239 CTTTTCATTCGTTCTCGTCACAATCCCTGGACCGAATAGCATTCATTTCGCCACGGGCA 20180

QY 179 CTAAATATAACAGATGATGACCTTATGACATATAGAGCTCGTAAGTGCCATCTAGAT 238

DB 20179 CCAACATAAGGGCATCCGGCATGCCCTTCTATACAGTGCCTCATAAAGGCGACCTACCAAT 20120

QY 239 GCCAGATTCGAACTGTTATTTAGGCGAACTCAACTAAGCGTAAATAATCCCTCTCAACT 298

DB 20119 ACTTCTTGTAACTGTTATTTATGTAACAACTCTATCAAGCTCAATGCTTCTCAGATT 20060

QY 299 ACCTTAGTAATAATCACATAGCTCCAAATCGTATCTCTCTAGTATATGAATACCTTCTC 358

DB 20059 TCCACCCCAATCTAAACACATGCCCCTTTGCAATATCTTCTAGAGTCCGAATGCTTGCCTC 20000

QY 359 AATTGACCATCGTCTGAGGATGAATG 387

Db 19999 GGATTGCCCATCGGTTTGATGATGGTACG 19971

RESULT 11

ADL17884/c
ID ADL17884 standard: cDNA: 158001 BP.

AC ADL17884;

DT 06-MAY-2004 (first entry)

DE Human phosphotyrosyl phosphatase activator, PTPA, gene.

Human; ds; antisense; phosphotyrosyl phosphatase activator; PRPA;
KW hyperproliferative disorder; developmental disorder; infection;
KW inflammation; tumour; gene.

Homo sapiens.

XX PN US2004023906-A1.

05-FEB-2004

XX DE 01-AUG-2002.

XX
BB
01 - NYC-2002. 2002US-00211178

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX

XXX

XX

PT phosphotyrosyl phosph

PT disorders.

PS Example 15; SEQ ID NO 11; 131pp; English.

The invention relates to a compound 8-80 nucleobases in length targeted to a nucleic acid molecule encoding phosphotyrosyl phosphatase activator (PTPA), that specifically hybridises with the nucleic acid molecule encoding phosphotyrosyl phosphatase activator and inhibits the expression of phosphotyrosyl phosphatase activator, i.e. an antisense oligonucleotide. Also included are a composition comprising the compound and a pharmaceutical carrier or diluent, a method of inhibiting the expression of phosphotyrosyl phosphatase activator in cells or tissues, a method of treating an animal having a disease or condition associated with phosphotyrosyl phosphatase activator and a method of screening for an antisense compound. The disease or condition is a hyperproliferative disorder or developmental disorder. The compound, particularly the antisense oligonucleotide is useful in modulating the function of nucleic acid molecules encoding phosphotyrosyl phosphatase activator. The antisense compound can also be used as research tools and diagnostics. It can also be used as tools in differential and/or combinatorial analyses to elucidate expression patterns of a portion or the entire complement of genes expressed within cells and tissues. The compound can also be used for treating diseases or conditions associated with phosphotyrosyl phosphatase activator, preferably hyperproliferative disorder or developmental disorder. The compound can also be used as prophylaxis, e.g. to prevent or delay infection, inflammation or tumour formation. The present sequence is the human PTPA gene which is a target for the antisense oligonucleotides of the invention.

Sequence 158001 BP: 34764 A: 40497 C: 41696 G: 40433 T: 0 U: 611 Other:

Query Match	7.3%	Score 104.8;	DB 12;	Length 158001;
Best Local Similarity	49.7%;	Pred. No. 4e-06;		
Matches 297;	Conservative	0;	Mismatches 297;	Indels 4; Gaps 1;

QY 500 AATTTTATTATTACCTCCGCTAGATTGGTAAATACTATTGCATTTATCTCATTTCAAT 559

[illegible]

RESULT 12

ABQ75107
ID ABO75107 standard; cDNA: 4985 BP.

XX
AC
ABQ75107:

XX DT 01-NOV-2002 (first entry)

XX DE bronhelles cambiae odourant receptor 2 genomic DNA SEQ ID NO:10.

XX Anopheles gambiae; mosquito; olfactory gene; arrestin 1; pest control;
KW odourant receptor; olfaction; gene; ds.
KW

XX Anonheles gambiae.

XX
DN
W030025927A-A2

XX
01-2003

XX
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
8

[illegible]

PR 24-JAN-2002; 2002US-00056405.

PA (UYVA-) UNIV VANDERBILT.

PI Zwiebel LJ;

DR WPI; 2002-627421/67.

XXXXXX

new mosquito larvicide, i.e. controlling the pest and disease vectors, on PT for identifying pest control agents.

Db 12973 GCAAACCTCTACCAGACTCAAGTGATCCGCCCAATGACCTCCCAATCCAAGACACACATC 13032
QY 324 CAAATCGTATCCCTCTAGTATATGAATCAACCTCTCAAAATTGACCATCGGTCTGAGGATG 383
Db 13033 CGCAACATATCCCTCAAGTGT-TGAATAGTCTCTCAGATTTGCCATTTGCTGATGAGGATGA 13091
QY 384 AATGCAG 390
Db 13092 TRAGCCG 13098

RESULT 15
ACF62794
ID ACF62794 standard; DNA; 8222 BP.
AC ACF62794;
AC ACF62794;
DT 09-OCT-2003 (first entry)
XX
DE Colon cancer analysis related genomic DNA SEQ ID NO:43.
XX
DE Human; colon cancer; oestrogen receptor; myoglobin; p21; p27; p16; p53;
KW progesterone receptor; pcna; cea; cdc2; c-erbB2; methylation; CpG;
KW characterisation; classification; diagnosis; differentiation;
KW colon cell proliferative disorder; gene; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO2003014388-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-EP008939.
XX
PR 09-AUG-2001; 2001DE-01039283.
XX
XX (BPIG-) EPIGENOMICS AG.
XX
PI Distler J, Model F, Taubert H;
XX
XX WPI; 2003-256600/25.
XX
PT Determining methylation status of CpG dinucleotides using modified
PT genomic sequences, oligonucleotides and/or PNA-oligomers, useful in the
PT characterization, grading, staging and/or diagnosis of colon cancer.
XX
XX Claim 22; Page 81-83; 219pp; English.

XX The present invention describes a method for determining the methylation
XX status of CpG dinucleotides within the genes for oestrogen receptor, p21,
XX p27, p16, progesterone receptor, myoglobin, pcna, cdc2, c-erbB2, p53
XX and/or cea, which comprises contacting the target nucleic acid with a
XX reagent that distinguishes between methylated and non-methylated CpG
XX dinucleotides, and determining from the methylation status of the CpG
XX positions the presence of a colon cancer. A set of oligomers or peptide
XX nucleic acid (PNA)-oligomers can be used as probes for determining the
XX cytosine methylation state and/or single nucleotide polymorphisms (SNP)
XX of a corresponding genomic DNA by analysis of a chemically pretreated
XX genomic DNA. The pretreated genomic DNA is useful for the determination
XX of the methylation status of a corresponding genomic DNA and/or detection
XX of SNPs. The methods and pretreated genomic DNA are also useful for the
XX characterisation, classification, diagnosis and differentiation of colon
XX cell proliferative disorders. ACF62752 to ACF63278 represent sequences
XX used in the exemplification of the present invention

XX
XX Sequence 8222 BP; 2010 A; 116 C; 1769 G; 4327 T; 0 U; 0 Other;
Query Match 7.2%; Score 103.4; DB 8; Length 8222;
Best Local Similarity 49.8%; Pred. No. 8.3e-06;
Matches 344; Conservative 0; Mismatches 341; Indels 6; Gaps 3;
QY 412 ATGGTACCTATAAAAAAATTATTATTTTAAAAAAATTGATGTGACCAGTGGTTGGAGAGA 471

Db 7223 ATGGTGTGTTTGTGTTTAAAGTTTGAATTAATAATGTTTATTGAGTAAATGGGTGGATG 7282
QY 472 GAGGTCTACCGATTGGTCAAGTGGCACCACAAATTTTATTATTTTACCTCCCTGAGATTCT 531
Db 7283 TTTGTAAGAAAGAGAGAGTGGATGATGTTAGTTTATTATTTATTTATTTATTTTATTT 7342
QY 532 AAATACTATTGCAATTTATCTCATTTTCAATTTTATTATTTTAAATTAATTTTATTTATTTGGA 589
Db 7343 AATTTATTAGGTATTAAATTTATTATTTATTATTTATTTTATTTAGTTATTATT 7402
QY 590 TAAAAATCTCAATACCTTTACTTTTTTTTAAAAAGAAATTTAATAATTTATTTATTTATTT 649
Db 7403 TATTTATTATTGTTGTTATTTTAAATTTATTATTTATTTTATTTTGTGTTATTTATTTAT 7462
QY 650 TAGATAAAAAATTTCTAATACTTTTACTTTTTTTTTTAAAAAGAAATTTCAATTCGCTTTTCT 709
Db 7463 TTTATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7522
QY 710 TAAATTTAGTTTAAATTTCTATACCTAATTTATAAAAAATTTCTGATCGGATTTAGTGGTCAA 769
Db 7523 CGATATTTTAAATTTATTATTTATTATTATTATTATTATTATTATTATTATTATTATTATT 7582
QY 770 AGTCAAGTCACATGAATTTTGTGGAGAAAAAATAAAAAATTAACACATTTTTCGATTAA 829
Db 7583 TTTATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7642
QY 830 TTTATTTATATATATAATAATAAACAACATTTTATTATTATTTTAAATGTTGCAATAATTTT 889
Db 7643 TTTATTTATTATTGTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7702
QY 890 AATTAAAA--ATTTTCAGCACACAAATTACACTCTCATTCATTAATAATTTAATTTTACCAT 947
Db 7703 TATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7762
QY 948 AATTAATAATTTGGAGACAAATTTTAAATCTCACCCCTCCATTAATGCAATTAATTAA 1007
Db 7763 TTTTATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7822
QY 1008 TTTTGTTCGA--TACTTCTTATTTCACCTCCCTAACATTAATCAATTAACCAATTTTGAAC 1065
Db 7823 TTTATTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7882
QY 1066 TGTATAAATTTCTTAACCTTTATTCACCTATTGT 1096
Db 7883 TATTTATTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7913

Search completed: August 13, 2005, 08:07:31
Job time : 1238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 12, 2005, 18:51:23 ; Search time 10337 seconds
(without alignments)
6717.261 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggacgcgtgttc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1433	100.0	1433	6	AX370650	Sequence
2	1393.8	97.3	5765	8	AF487511	Populus b
3	266	18.6	572	8	AF025353	Gossypium
4	184	12.8	300	8	AF060583	Gossypium
5	179.4	12.5	5998	8	AF060583	Gossypium
6	179	12.5	913	8	AY181254	Gossypium
7	156.4	10.9	927	8	AF276939	Gossypium
8	150.6	10.5	111063	2	AC146584	Medicago
9	140.2	9.8	121257	2	AC146307	Medicago
10	139.8	9.8	119270	2	AY508219	Medicago
11	139	9.7	96930	8	AC149484	Populus b
12	137.6	9.6	78069	2	AC150760	Medicago
13	137.4	9.6	141416	8	AC149480	Populus b
14	137.2	9.6	143392	8	AC149299	Populus b
15	136.4	9.5	107287	2	AC146794	Medicago
16	136.4	9.5	181249	2	AC148479	Zea mays
17	135.6	9.5	65127	2	AC148403	Medicago
18	135.6	9.5	123241	2	AY379775	Medicago
19	135.4	9.4	106114	2	AC146565	Medicago

C 20	135.4	9.4	122167	8	AC146806	Medicago
C 21	135	9.4	105035	2	AC145452	Zea mays
C 22	134	9.4	170017	2	AC148158	Zea mays
C 23	133.8	9.3	137841	2	AC147714	Medicago
C 24	133.6	9.3	100595	8	AP004545	Lotus cor
C 25	133.4	9.3	116840	2	AC141862	Medicago
C 26	133.4	9.3	125529	2	AC121242	Medicago
C 27	132.4	9.2	116004	8	AC149208	Medicago
C 28	132.4	9.2	135726	8	AC147498	Medicago
C 29	132.2	9.2	125478	8	AC147430	Medicago
C 30	132	9.2	125522	2	AC148718	Medicago
C 31	132	9.2	198102	2	AC145389	Zea mays
C 32	131.8	9.2	86095	8	AC136450	Medicago
C 33	131.2	9.2	113049	8	AC146649	Medicago
C 34	131	9.1	185534	2	AC145481	Zea mays
C 35	130.6	9.1	78128	8	AC144805	Medicago
C 36	130.6	9.1	119996	2	AC137666	Medicago
C 37	130.4	9.1	171261	2	AC146349	Zea mays
C 38	130.4	9.1	178675	2	AC145262	Medicago
C 39	130.2	9.1	44109	2	AC146805	Medicago
C 40	130.2	9.1	110625	2	AC138199	Medicago
C 41	130.2	9.1	124889	8	AC126008	Medicago
C 42	130.2	9.1	133526	2	AC135315	Medicago
C 43	130.2	9.1	137686	2	AC130806	Medicago
C 44	130	9.1	206462	2	AC148173	Zea mays
C 45	129.6	9.0	132211	8	AC135396	Medicago

ALIGNMENTS

RESULT 1
AX370650 LOCUS AX370650 1433 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 2 from Patent WO0210377.
ACCESSION AX370650
VERSION AX370650.1 GI:19168815
KEYWORDS
SOURCE Arabidopsis sp.

ORGANISM Arabidopsis sp.
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1
AUTHORS Cai, L., Li, X., Cheng, N. and Liu, J. W.

TITLE Isolation and characterization of a fiber-specific g(b)-tubulin

JOURNAL Promoter from cotton

PATENT Patent: WO 0210377-A 2 07-FEB-2002;

Institute of Molecular Agrobiolgy (SG)

FEATURES
Location/Qualifiers

source 1..1433

/organism="Arabidopsis sp."

/mol_type="unassigned DNA"

/db_xref="taxon:29726"

ORIGIN

Query Match 100.0%; Score 1433; DB 6; Length 1433;

Best Local Similarity 100.0%; Pred. No. 7.4e-189;

Matches 1433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTATAGGCACGCGTGTTCAGACGCGCGGCGTGTTCAGATCTATGATATTTTCAGAT 60

Db 1 ACTATAGGCACGCGTGTTCAGACGCGCGGCGTGTTCAGATCTATGATATTTTCAGAT 60

Qy 61 TTGCATTAAGACTTCTATCTATCAGAACGCGCTCAGAGGATCCCAATTTAGTCTAAAT 120

Db 61 TTGCATTAAGACTTCTATCTATCAGAACGCGCTCAGAGGATCCCAATTTAGTCTAAAT 120

Qy 121 TATCTTCAGTCTCGGAACCAACTCAGACCCCAAAACCCGCGCTCAGCCCACTCAGTCT 180

Db 121 TATCTTCAGTCTCGGAACCAACTCAGACCCCAAAACCCGCGCTCAGCCCACTCAGTCT 180

Qy 181 AATATAACAGATATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATGC 240

181 AATATAACAGATGATGACCTTATGACCATATAGAGCCTCGTAGGTCGCATCTAGATGC 240
241 CAGATTGGAACCTGTTATTGAGCGAAGCTCACTAACGCTGAAGAAATCTCTCAACTAC 300
241 CAGATTGGAACCTGTTATTGAGCGAAGCTCACTAACGCTGAAGAAATCTCTCAACTAC 300
301 CTTAGTAATAAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCTTCTCAA 360
301 CTTAGTAATAAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCTTCTCAA 360
361 ATTGACCATCGGTCGAGAGTGAATGACGACCGGTGCGCCACCGATTTACTTAATGTTACCT 420
361 ATTGACCATCGGTCGAGAGTGAATGACGACCGGTGCGCCACCGATTTACTTAATGTTACCT 420
421 ATAAAAAATTAATTTATTTTAAAAAATGATGACGAGTGGTGGAGAGAGGTTCTAC 480
421 ATAAAAAATTAATTTATTTTAAAAAATGATGACGAGTGGTGGAGAGAGGTTCTAC 480
481 CGATTGCTCAAGTGGACCAATTTTATTTTATTTTACCTCCTGCTAGATTCTGAATACTAT 540
481 CGATTGCTCAAGTGGACCAATTTTATTTTATTTTACCTCCTGCTAGATTCTGAATACTAT 540
541 TGCATTTATCTCAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
541 TGCATTTATCTCAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
601 ATACTTTACTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
601 ATACTTTACTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
661 TCTAATCTTTACTTTTAAAAAGAAATTTCAATTTGCTGTTTCTTAAATTTAGTTT 720
661 TCTAATCTTTACTTTTAAAAAGAAATTTCAATTTGCTGTTTCTTAAATTTAGTTT 720
721 TAATTTCTACTAATTAATAAAATCTGATCGGAAATGATGCTGCTGCTGCTGCTGCTGCTGCT 780
721 TAATTTCTACTAATTAATAAAATCTGATCGGAAATGATGCTGCTGCTGCTGCTGCTGCTGCT 780
781 ATGAATTTGCTGAGAAAAATTAATAAAATTAATAAAATTTTCAATTTATTTATTTATTTATTA 840
781 ATGAATTTGCTGAGAAAAATTAATAAAATTAATAAAATTTTCAATTTATTTATTTATTTATTA 840
841 TATAAATAATAAACAATTTTATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
841 TATAAATAATAAACAATTTTATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
901 CAGCAACAATTAACATCTCACTCAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 960
901 CAGCAACAATTAACATCTCACTCAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 960
961 AGGCAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1020
961 AGGCAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1020
1021 CTTCTTATTTTCACTTCAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTCTTA 1080
1021 CTTCTTATTTTCACTTCAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTCTTA 1080
1081 ACTTATTTCACTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1081 ACTTATTTCACTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1141 CACACTTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 CACACTTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 AATTTTCAAAATTTCAATTTTCCAGGAAAAAGAAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1201 AATTTTCAAAATTTCAATTTTCCAGGAAAAAGAAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1261 CAGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

1261 CAGCTCAACCTGCGCTCAATTAACGGAATCTTTTAACTCCTCTATATACCCAAACCA 1320
1321 CTCTCATACCAATTTCCCATAAAAAAGATTTCCGGAATCTTTATTTCTTTTATATTTT 1380
1321 CTCTCATACCAATTTCCCATAAAAAAGATTTCCGGAATCTTTATTTCTTTTATATTTT 1380
1381 CCTCTCAATTTCCCTGCTCACTTTCCGAGAAAAATGAGAGAAATCTTTCATC 1433
1381 CCTCTCAATTTCCCTGCTCACTTTCCGAGAAAAATGAGAGAAATCTTTCATC 1433

RESULT 2
AF487511 5765 bp DNA linear PLN 17-OCT-2002
LOCUS Gossypium hirsutum beta-tubulin (TUB1) gene, complete cds.
DEFINITION AF487511
ACCESSION AF487511
VERSION AF487511.1 GI:19569608
KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 5765)
Li, X.B., Cai, L., Cheng, N.H. and Liu, J.W.
Molecular Characterization of the Cotton GhTUB1 Gene That Is
Preferentially Expressed in Fiber
Plant Physiol. 130 (2), 666-674 (2002)
22263995
MEDLINE 12376634
PUBMED
REFERENCE 2 (bases 1 to 5765)
Li, X.B., Cai, L. and Liu, J.W.
Direct Submission
Submitted (25-FEB-2002) Laboratory of Plant Reproduction and Cotton
Biotechnology, Institute of Molecular Agrobiolgy, 1 Research Link,
NUS, Singapore 117604, Republic of Singapore
LOCATION/Qualifiers
1. 5765
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"
1377..>3349
/gene="TUB1"
join(1377..1771,1876..2144,2676..>3349)
/product="beta-tubulin"
join(1377..1771,1876..2144,2676..3349)
/gene="TUB1"
/note="microtubule protein subunit"
/codon_start=1
/product="beta-tubulin"
/protein_id="AAL92118.1"
/db_xref="GI:19569609"
/translation="MREILHIOGGCGNOIGAKFWEVVCVCAEHGIDSTGRYGGSELOL
ERLNVINEASCRFVPRVAVLMLEPGTWDVSRSGYQIIPRPNFVRGSGAGNWA
KHYTGAEILIDVLDVVRKEAENCLQGFQVCHSLGGTSGMGTLILSKIRBEP
DRMLTFVFPSPKSDTVVEPNATLSVHQLVENADECMVLDNEALYDLCFTLKUT
TPSGDLNHLISATMSGVTCCLFPQGLNSDLKLAIVNLIFFPLHFFMVFAPLTSR
GSQOYRALTVPELTQOMDKQNMCAADPRHGRYLTASAVRGMSTKEVDEQMINVQ
KNSSYFVSEWIPNNVKTVDIPPIGLKMASTFGNSTSIQEMPRRVSEQFTAMFRK
AFUHWTTGEGMDEMEFTEAESNMNDLVSEYQQYQDATDDEEYEEBEYEAEA"

Query Match 97.3%; Score 1393.8; DB 8; Length 5765;
Best Local Similarity 99.9%; Pred. No. 1.4e-183;
Matches 1395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

37 GCTTGATATCTATGATTTTCAGATTTGCAATGCACTTCTATCTATCAGAAGACGCTGCA 96
1 GCTTGATATCTATGATTTTCAGATTTGCAATGCACTTCTATCTATCAGAAGACGCTGCA 60
97 GAGGATCCCAAAATAGTCTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAA 156

Db 61 GAGGATCCCAATTAGTCTAAAAATATCTTCAGTCTCGGAAACCAACTCAGGACCCAAAA 120
Qy 157 CCCGTCGCTCACCCCACTCAGTCTAAATATAACAGAGTATGACACTTATGACCATATAGAG 216
Db 121 CCCGTCGCTCACCCCACTCAGTCTAAATATAACAGAGTATGACACTTATGACCATATAGAG 180
Qy 217 CCTCGTAAGGTGCCATCTAGATGCCAGATCGGAACTGTTATTTGAGCGAACTCAACTA 276
Db 181 CCTCGTAAGGTGCCATCTAGATGCCAGATCGGAACTGTTATTTGAGCGAACTCAACTA 240
Qy 277 ACGTAAAAATCTCTCAACTACCTAGTATTAATAACATACATAGTCCAAATCGTATCCT 336
Db 241 ACGTAAAAATCTCTCAACTACCTAGTATTAATAACATACATAGTCCAAATCGTATCCT 300
Qy 337 CTAGTATATGAATCAACCTTCTCAAAATGACCATCGTCTGAGGATGGAATGCAGACCGGT 396
Db 301 CTAGTATATGAATCAACCTTCTCAAAATGACCATCGTCTGAGGATGGAATGCAGACCGGT 360
Qy 397 GCCACCGATTTACTAATGGTACCTATAAAAAATTAATTAATTTTAAAAAATTTGATGTGAC 456
Db 361 GCCACCGATTTACTAATGGTACCTATAAAAAATTAATTAATTTTAAAAAATTTGATGTGAC 420
Qy 457 CAGTGGTTGGAGAGAGGCTCTACCGATTGCTCAAGTGGCACCAATTTTATTTTACCT 516
Db 421 CAGTGGTTGGAGAGAGGCTCTACCGATTGCTCAAGTGGCACCAATTTTATTTTACCT 480
Qy 517 CCTGCTAGATTCGTAAATACCTATTGCAATTTATCTCAATTTATTTTAAATTAATTT 576
Db 481 CCTGCTAGATTCGTAAATACCTATTGCAATTTATCTCAATTTATTTTAAATTAATTT 540
Qy 577 TATATTAATTTGGATAAAAAATCTAATACTTTTACTTTTTTAAAAAAGAAATTTATTAAT 636
Db 541 TATATTAATTTGGATAAAAAATCTAATACTTTTACTTTTTTAAAAAAGAAATTTATTAAT 600
Qy 637 ATTTTATTAATTTAGATAAAAAATCTAATACTTTTACTTTTTTAAAAAAGAAATTTCA 696
Db 601 ATTTTATTAATTTAGATAAAAAATCTAATACTTTTACTTTTTTAAAAAAGAAATTTCA 660
Qy 697 TTGGGTTTTTTCTTAATTTAGTCTTAATTTCTATTAATAAATAATTTCTGATCGGAT 756
Db 661 TTGGGTTTTTTCTTAATTTAGTCTTAATTTCTATTAATAAATAATTTCTGATCGGAT 720
Qy 757 AGTGTGGTGTCAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATAAATAATTTAAACAC 816
Db 721 AGTGTGGTGTCAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATAAATAATTTAAACAC 780
Qy 817 ATTTTTCGATTAATTTAT 876
Db 781 ATTTTTCGATTAATTTAT 840
Qy 877 AATAATATTTTAAATTAATAATTTTCAACACAACTATCACTCTCATCAATTAATTTAAT 936
Db 841 AATAATATTTTAAATTAATAATTTTCAACACAACTATCACTCTCATCAATTAATTTAAT 900
Qy 937 CTATTTACCAATAATTTAAATTTGTGAGGCAATTAATTTTAAATCTCACCTCCCAATTAATG 996
Db 901 CTATTTACCAATAATTTAAATTTGTGAGGCAATTAATTTTAAATCTCACCTCCCAATTAATG 960
Qy 997 CATATTAATTAATTTTGTTCGACTTCTTATTTTCACTCTCTCAACATTAATCAATTAACCCA 1056
Db 961 CATATTAATTAATTTTGTTCGACTTCTTATTTTCACTCTCTCAACATTAATCAATTAACCCA 1020
Qy 1057 ATTTTGAAGTGTATTAATTTCTTAATTTTCACTATTTGTCGCTCTGGGTCATCTGAA 1116
Db 1021 ATTTTGAAGTGTATTAATTTCTTAATTTTCACTATTTGTCGCTCTGGGTCATCTGAA 1080
Qy 1117 AGGCCACCGTCTCAGGCTGTCAAACCACTTTTGCACGTCTCAATTTCCAGTAACTACAT 1176
Db 1081 AGGCCACCGTCTCAGGCTGTCAAACCACTTTTGCACGTCTCAATTTCCAGTAACTACAT 1140
Qy 1177 TGTTACAGTTACTAAGCAAAATCCCAATTTTCAAAAATTTCAAAATTTCCAGGAAACGAAACG 1236
Db 1141 TGTTACAGTTACTAAGCAAAATCCCAATTTTCAAAAATTTCAAAATTTCCAGGAAACGAAACG 1200

Qy 1237 TCCGTTACTAACCGACCTAAACCCAGCTCAACCTGCCGTCAATTAACGGAATCTTTTA 1296
Db 1201 TCCGTTACTAACCGACCTAAACCCAGCTCAACCTGCCGTCAATTAACGGAATCTTTTA 1260
Qy 1297 ACTCCTCTATATAACCAACCACTCTCATCACCATTTCCTCAATAAAGAAATTTCCGG 1356
Db 1261 ACTCCTCTATATAACCAACCACTCTCATCACCATTTCCTCAATAAAGAAATTTCCGG 1320
Qy 1357 AATTTCTTATTTCTTTTATATTTTCTCTCCAAATTTCCCGTCACTTTCCGAGAAAATGA 1416
Db 1321 AATTTCTTATTTCTTTTATATTTTCTCTCCAAATTTTCCGTCACCTTTCCCGAGAAAATGA 1380
Qy 1417 GAGAAATCTTTTACATC 1433
Db 1381 GAGAAATCTTTTACATC 1397

RESULT 3
AF025353/c 572 bp DNA linear PLN 08-JAN-1999
LOCUS Gossypium hirsutum tandem repeat B77 sequence.
DEFINITION AF025353
ACCESSION AF025353
VERSION AF025353.1 GI:4128202

KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 572)
AUTHORS Zhao X., Ji Y., Ding X., Stelly D.M. and Paterson A.H.
TITLE Macromolecular organization and genetic mapping of a rapidly evolving chromosome-specific tandem repeat family (B77) in cotton (Gossypium)

JOURNAL Plant Mol. Biol. 38 (6), 1031-1042 (1998)
MEDLINE 93084754
PUBMED 9869409

REFERENCE 2 (bases 1 to 572)
AUTHORS Zhao X., Ji Y., Ding X., Stelly D.M. and Paterson A.H.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-1997) Dermatology, University of Michigan, 1500 E. Medical Center Dr., 3420 CCGC, Box 0932, Ann Arbor, MI 48109-0932, USA

FEATURES
Location/Qualifiers
source 1..572
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"

repeat_region 1..572
/xpt_family="B77"
/rpt_type=tandem

ORIGIN

Query Match 18.6%; Score 266; DB 8; Length 572;
Best Local Similarity 91.6%; Pred. No. 1.4e-27;
Matches 326; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

Qy 35 GTGCTTGATATCTAAGATTTTTCAGATTGCTAAGACTTCTATCTATCAGAGAGCGCTG 94
Db 421 GTACTTGAATCTCTGA-TTTTCAGATATGCTAAGACTTCTATCTATCAGAGAGCGCTT 363
Qy 95 CAGAGGATCCCAATTAAGTCTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCCAA 154
Db 362 CAC--GATCCCAATTAAGTCTAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCTGA 306
Qy 155 AACCCGTCGCTCACCCCACTCAGTCTAATAATAACAGAGTATGACATTTATGACCATATAG 214
Db 305 AACCCGTCGCTCACCCCACTCAGTCTAATAATAACAGAGTATGACATTTATGACCATATAG 247
Qy 215 AGCCTCTGAAGGTGCCATCTAGATGCCAGATTTGGAAGAACTGTTATTTAGGGGAACCTCAAC 274
Db 246 AACCTCTGAAGGTGCCATCTAGATGCCAGATTTGGAAGAACTGTTATTTAGGGGAACCTCAAC 187


```

REFERENCE
AUTHORS
  Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
  and Roe, B.A.
TITLE
JOURNAL
  Direct Submission
  Submitted (17-SEP-2004) Department of Chemistry And Biochemistry,
  The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
  OK 73019, USA
COMMENT
  On Sep 17, 2004 this sequence version replaced gi:52000565.
  ----- Genome Center
  Center: Department of Chemistry And Biochemistry
  The University Of Oklahoma
  Center code:UOKNOR
  -----
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 6 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submitter.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  *
  1 55923: contig of 55923 bp in length
  * 55924 56023: gap of unknown length
  * 56024 87161: contig of 31138 bp in length
  * 87162 87261: gap of unknown length
  * 87262 92041: contig of 4780 bp in length
  * 92042 92141: gap of unknown length
  * 92142 105791: contig of 13650 bp in length
  * 105792 105891: gap of unknown length
  * 105892 108659: contig of 2768 bp in length
  * 108660 108759: gap of unknown length
  * 108760 111063: contig of 2304 bp in length.
  Location/Qualifiers
    1. 111063
      /organism="Medicago truncatula"
      /mol type="genomic DNA"
      /db xref="taxon:3880"
      /clone="mth2-68k18"
      /clone_lib="Medicago truncatula BAC library H2"
  ORIGIN
    Query Match 10.5%; Score 150.6; DB 2; Length 111063;
    Best Local Similarity 49.9%; Pred. No. 4.1e-12;
    Matches 440; Conservative 0; Mismatches 429; Indels 12; Gaps 2;
  Qy 27 CCGGGCTGGTCTGATATCATGATTTTCAGATTTCGATTAAGACTTCTATCTATCAAGAA 86
  Db 41044 CCTTCTGAAACTCAAGATCTTTCTACGCTTATCATGATAAATCTTTTGTGCGACTCTGC 41103
  Qy 87 GAGCCTCGAGAGATCCCAATTAGTCTAAATATCTTCAAGTCTCGAACCAGCACTCA 146
  Db 41104 GACGCTTTCATTTCTCTCAGATCATCTGAACCTTCTCAGTAGTTTGTGGAACAATCTCT 41163
  Qy 147 GGACCCAAACCGCTGCTCACCAGTCTAGTCTAATATACAGAGATGACACTTATGA 206
  Db 41164 GGTCTAAGACCATCTTTCACCTGACTCAACACGACAGGAGTCTGCTCCTCGA 41223
  Qy 207 CCATATAGAGCCTCGTAGGTGCCATCTAGATGCCAGATTGGAACATGTTATTGTAGCG 266
  Db 41224 CCATATAAGCCTCGAAGGTGCCATCTCAATCTGGAATGATAGCTATTATTGTATGTG 41283
  Qy 267 AACTCAACTACGCTTAAATCTCTCACTACTTATAGTAATAAATACATAGCTCTCAA 326
  Db 41284 AACTCGATCAACGGAAGATGACTATCCCAAGTCTCTCTGCTCAAGAACAAATCCTC 41343
  Qy 327 ATCGTATCCTCTAGTATATGATCACTTCTCAAAATGACCATCGGCTCTGAGGTGGAAT 386
  Db 41344 AACAAATCCTCTAGGACTGAATGTTCTCTCCGACTGACCATCTATCTGTGGATGATAC 41403
  Qy 387 GCAGACGGTGCACCGGATTACTTAATGGTACCTATATAAATATTATTTTAAAAAA 446

```

```

Db 41404 GCCGAACCTCAATCTCAACTTTCACATAATGTATCAAGTGTGAATGAAGTCTTTAAAAAAA 41463
Qy 447 TTGATGTGACCACTGGTGTGGAGAGAGAGGTCTACCGATTGGTCAAGTGGCACCACAAATTTT 506
Db 41464 TGTGAATTAAGATTATATATTTGGATGAATAGTAGAATTTGAACA-----ACAAATAAGT 41518
Qy 507 TATTTTACCTCCGCGCCAGATTGCGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 566
Db 41519 GAGATGACCCATATAACTAATTTCTTAAGATTTTAGGGTGACATGCGATATCAAACTCGAT 41578
Qy 567 TTAATTTATTTATTTATTTTGGATAAAATTTCTAATACCTTTTCTTTTAAAAAGAAT 626
Db 41579 GTAGGATACCGAAGAGTTCCAAAATATATGATGAATATTTGTTAGTTTATAGATATCA 41638
Qy 627 TTAATTTAATTTATTTATTTATTTTAGATAAAATTTCTAATACCTTTTCTTT-----TTT 679
Db 41639 TTTTAAATTTTTCATCTTAATTTTGTATTTGGATGTCATTTCACTTTTCATGCTATTT 41698
Qy 680 TTTAAAAAGAAATTTCAATTTGCGTTTTTCTTAATTTAGTTTAAATTTCTATATCAATTAATA 739
Db 41699 TTTTACTCAACCTCCATTTTTCACGCTTTTGTGAAATTTTAAATGACGTGACGTTT 41758
Qy 740 AAAATTTCTGATCGATTAGTGTCTCAAGTCAAGTCAATGACATGAAATTTTGTGGAGAAA 799
Db 41759 TTAATTTGAGCCTAATTTTATATAAAATATATCTCAAAATCTTCAATCAATCAACAA 41818
Qy 800 AAATAAAAAATTAACACATTTTTCGATTAATTTATTTATATATATATAATATAATATAACACAT 859
Db 41819 AAAATTAACAAAAATTAATCACTCACTTTCTATATATCTCGGTTAAACACAA 41878
Qy 860 TTTTATTTAATGTTGTCATAATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Db 41879 TTTTCTTCTATCACTTTCTCATGTTAAATTTGATTTT 41919
RESULT 9
AC146307 121257 bp DNA linear HTG 09-SEP-2004
LOCUS Medicago truncatula clone mth2-9p17, WORKING DRAFT SEQUENCE, 3
DEFINITION ordered pieces.
ACCESSION AC146307 GI:51948617
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Medicago truncatula BAC Clone mth2-9p17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2004) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Sep 9, 2004 this sequence version replaced gi:51699618.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

```



```
VERSION AC149484.1 GI:48209803
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 96930)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 96930)
AUTHORS Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
CONSTRM DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated total Number of Errors is 0.1.
FEATURES
    source
    1..96930
        /organism="Populus balsamifera subsp. trichocarpa"
        /mol_type="genomic DNA"
        /sub_species="trichocarpa"
        /db_xref="taxon:3694"
        /clone="Popl-69M19"
..
ORIGIN
    Query Match 9.7%; Score 139; DB 8; Length 96930;
    Best Local Similarity 63.1%; Pred. No. 1.7e-10;
    Matches 214; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
Qy 52 TTTTCAGATTGCATAAGACTTCTATCTATCAGAGACGCGTGCAGAGATGCCAAATTA 111
Db 54757 TTTCTTTATCTGCGTAACCTCTTTTGTCTGCTGAGCTGTTGAAGCTTATTCTTAATTA 54816
Qy 112 GTCATAAATTCCTTCAGTCTCGGAACCACTCAGGACCCAAACCGCTCGTCCACCA 171
Db 54817 CCTCTATCTTCTGAGGTAATCTGAATCAACTCTGCTGCCATAGCCTCTTCTCACCA 54876
Qy 172 ACTCAGTCTAATATAACAGAGTATGACATTTATCACCATATAGAGCCTCGTAAGGTGCCA 231
Db 54877 CCTCAAACCAACAAACCGGTGATCTACATCTCCAGCATTACAAAGCTCATAGGTGCCA 54936
Qy 232 TCTAGATGCCAGATTGGAACTGTTATTTGTAGGCGAACTCACTAAACGGTAAATAATCCT 291
Db 54937 TCTCTATGCTAGCCTGATACCTGTTGTTGAAGCAAAATTCACAAATGGTAAGAACTTAC 54996
Qy 292 CTCAACTACCTTAGTAATAATACATAGCTCCAAATCGTATCTCTAGTATATGATCA 351
Db 54997 TCCAAACCAACTCCAAAATCCATAACACAGGCCCTTAGCATATCTCCAAGATCTGAATAG 55056
Qy 352 CTTTCTCAAAATTGACCATCGGTCTGAGGATGGAATGCAG 390
Db 55057 TCCTCTAGATTGACCATCTGCTGAGGGTGAAAAGCTG 55095
RESULT 12
LOCUS AC150760/c 78069 bp DNA linear HTG 10-OCT-2004
DEFINITION Medicago truncatula clone mth2-171i12, WORKING DRAFT SEQUENCE, 15
unordered pieces.
ACCESSION AC150760
VERSION AC150760.3 GI:54020976
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Medicago truncatula BAC Clone mth2-171i12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Oct 10, 2004 this sequence version replaced gi:54019504.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3217: contig of 3217 bp in length
* 3218 3317: gap of unknown length
* 3318 5758: contig of 2441 bp in length
* 5758 5858: gap of unknown length
* 5859 8001: contig of 2143 bp in length
* 8002 8101: gap of unknown length
* 8102 11395: contig of 3294 bp in length
* 11396 11495: gap of unknown length
* 11496 14844: contig of 3349 bp in length
* 14845 14944: gap of unknown length
* 14945 18915: contig of 3971 bp in length
* 18916 19015: gap of unknown length
* 19016 23663: contig of 4648 bp in length
* 23664 23763: gap of unknown length
* 23764 28054: contig of 4291 bp in length
* 28055 28154: gap of unknown length
* 28155 33934: contig of 5780 bp in length
* 33935 34034: gap of unknown length
* 34035 38181: contig of 4147 bp in length
* 38182 38281: gap of unknown length
* 38282 44430: contig of 6149 bp in length
* 44431 44530: gap of unknown length
* 44531 50090: contig of 5560 bp in length
* 50091 50190: gap of unknown length
* 50191 57264: contig of 7074 bp in length
* 57265 57364: gap of unknown length
* 57365 64151: contig of 6787 bp in length
* 64152 78069: gap of unknown length
* 78069 84252: contig of 13818 bp in length.
    Location/Qualifiers
    1..78069
        /organism="Medicago truncatula"
        /mol_type="genomic DNA"
        /db_xref="taxon:3880"
        /clone="mth2-171i12"
        /clone_lib="Medicago truncatula BAC library H2"
ORIGIN
```

Query Match 9.6%; Score 137.6; DB 2; Length 78069;
Best Local Similarity 57.8%; Pred. No. 2.8e-10;
Matches 245; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 27 CCGGGCTGGCTGGATATCATGATTTTCAGATTTCGATTAAGACTTCTATCATACGAA 86
Db 29516 CTTTCCTGAAACTCAAGATCTTTTCTAGCTTATCATGATAACTCTTTTGTGCAATCTGC 29457

Qy 87 GAGCCCTGCAGAGGATCCAAATTTAGTCTAAATTTATCTTCAGTCTCGGAACCAACTCA 146
Db 29456 GAGCCCTTCATTTCTCTTGATGATCATCGAACTTTCTCAGTAGTCTGCTGACAACTCTCT 29397

Qy 147 GGACCCAAACCCGCTGGCTACCCCAACTCAGTCTTAATATACAGAGATGACACTTATGA 206
Db 29396 GATCCTAAGACCACTCTTTACCTCACTCAAACCCAGCACAGGAGTCTGCACTCCGA 29337

Qy 207 CCATATAGAGCTCGTAAGTGCCATCTAGATGCGAGATTGGAACCTGTTATTGTAGCG 266
Db 29336 CCATATAAGCTCGAAAGGTGCCATTTCCAACTACTAGAAATGATAACTATTATTGTATGTG 29277

Qy 267 AACTCACTAAGCGTAAATAATCTCTCAACTACCTTAGTAATAATAACACATAGCTCAA 326
Db 29276 AACTCGATCAACGAAGATGACTATCCCAAGTCTCTCTGCTCAAGACACAAATTTCTC 29217

Qy 327 ATCGTATCCTCTAGTATATGAATCACCTTTCTCAAATTGACCATCGGTCTGAGGATGGAAT 386
Db 29216 AACAAATCCTCTAGGACTGAATTTGCTCTCCAACTGACCATCTGCTCGGATGATAC 29157

Qy 387 GCAGACGGTCCACCGAATTTACTTAATGGTACTATATAAAATTTATTTTAAAAAA 446
Db 29156 GCCGAATCTCAACTCAACTTCGAACCCCAAGCCTTTGCAAACTCTTCCAAATCTAGAA 29097

Qy 447 TTGA 450
Db 29096 TTA 29093

RESULT 13
AC149480/c
LOCUS AC149480 141416 bp DNA linear PLN 04-JUN-2004
DEFINITION Populus balsamifera subsp. trichocarpa clone Popl-048004, complete sequence.
ACCESSION AC149480
VERSION AC149480.1 GI:48209775
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 141416)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Unpublished
JOURNAL 2 (bases 1 to 141416)
REFERENCE 2 (bases 1 to 141416)
AUTHORS Stanford Human Genome Center.
CONSTRM DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source 1. 141416
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="genomic DNA"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone="Popl-048004"

ORIGIN
Query Match 9.6%; Score 137.4; DB 8; Length 141416;
Best Local Similarity 62.8%; Pred. No. 2.6e-10;
Matches 213; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Qy 52 TTTTCAGATTGGATTAAGACTTCTATCTATCATCAGAAGCGCTGAGAGGATGCCAAATTA 111
Db 108532 TTTTATCTCGGTAACTCTTTTGTCTAGTGTGTTGAAGCTTCTGTCTAATTA 108473

Qy 112 GTCTAAATTTATCTTCAGTCTCGGAACCAACTCAGGACCCAAACCGCTGCCTCACCCA 171
Db 108472 CCTCTATCTCTCTGAGGTAATCTGAATAAACTCCGGTCCCATAGGCTCTTCTCACCA 108413

Qy 172 ACTCAGTCTAATAATAACAGAGATATGACACTTATGACCATATAGAGCTTCGTAAGGTGCCA 231
Db 108412 CCTCAACCAACAAACCGGTGATCTACACTTCGACCATACAAAGCTCTATAAGGTGCCA 108353

Qy 232 TCTAGATGCCAGATTTGGAACCTGTTATTGTAGGGGAACCTCAACTAACCGGTAAATAATCCT 291
Db 108352 TCTCTATGCTAGCTGATAACTGTTGTTGAAGCAAAATTCACCTAATGGTAGGAACCTTAC 108293

Qy 292 CTCACCTACCTTAGTATATAATCACATAGCTCCAAATCGTATCTCTAGTATATATATCA 351
Db 108292 TCCAAACCAACTCCAAATTCATGACACACAGCCCTCAACATATCTCCAAGATCTGAATAG 108233

Qy 352 CCTTCTCAAAATTGACCATCGTCTGAGGATGGAATGCAG 390
Db 108232 TCCTCTGAGTACCATCTGCTGAGGGTGAAAAGCTG 108194

RESULT 14
AC149299
LOCUS AC149299 143392 bp DNA linear PLN 02-JUN-2004
DEFINITION Populus balsamifera subsp. trichocarpa clone Popl-037B22, complete sequence.
ACCESSION AC149299
VERSION AC149299.1 GI:47900639
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 143392)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Unpublished
JOURNAL 2 (bases 1 to 143392)
REFERENCE 2 (bases 1 to 143392)
AUTHORS Stanford Human Genome Center.
CONSTRM DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source 1. 143392
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="genomic DNA"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone="Popl-037B22"

ORIGIN
Query Match 9.6%; Score 137.2; DB 8; Length 143392;
Best Local Similarity 62.0%; Pred. No. 2.8e-10;
Matches 217; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
Qy 38 CTTGATATCTATGATTTTCAGATTTGCATAGAATTTCTATCTATCAGAGACGCCCTGCAG 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108298 CTCTAGTCAGCATCTCTCTATCTGGTAATCTTTTGTCTACTCTGAGCTGTTTGAAG 108357
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 98 AGGATCCCAATTAAGTCTAAATTAATCTCAGTCTCGGAACCAACTCAGGACCCAAAC 157
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108358 CTTCTTTCTAATACCTCTATCTCTCTGAGGTAATCGAATCAACTCTGCTGCCATTAG 108417
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 158 CCGTCGCTACCCCACTCAGCTTAATATACAGAGTATGACACTTATGACCATATAGAGC 217
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108418 CCTCTCTACCAACCTCAACCAACCAACACCGGATCTACCTTCGGACCATACAAAGC 108477
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 218 CTGCTAGGTGCCATCTAGATGCGAGATTTGGAATCTGTTATTTAGGCGCAACTCAACTAA 277
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108478 CTCATAGGTGCCATCTTCTATGCTAGCTGATTAATCTGTTGTTGTAAGCAAAATCCACTAA 108537
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 278 CGGTAAAAATCTCTCAACTACTCTTAGTATTAATCAATCAGATGCTCAATCTGATCTCTC 337
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108538 TGGTAGGAATTAATCTCAACCAACTCCAAATCCATAACAAGCCCTTAACATATCTTT 108597
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 338 TAGTATATGAATCACCTTTCTCAATTCGACCATCGGTCTGAGGATGGAATG 387
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108598 CAAGATCTGAATAGTCTCTCAGACTGACCATCGGTCTGAGGATGGAAG 108647
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
AC146794/4
LOCUS AC146794 107287 bp DNA linear HTG 14-SEP-2004
DEFINITION Medicago truncatula clone mth2-1019, WORKING DRAFT SEQUENCE, 11
unordered pieces.
ACCESSION AC146794
VERSION AC146794.18 GI:50345165
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (Barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Medicago truncatula BAC Clone mth2-1019
Unpublished
REFERENCE 2 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (07-OCT-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (14-SEP-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Jul 16, 2004 this sequence version replaced gi:50057955.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
```

* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2068: contig of 2068 bp in length

```
* 2069 2168: gap of unknown length
* 2169 4879: contig of 2711 bp in length
* 4880 4979: gap of unknown length
* 4980 7400: contig of 2421 bp in length
* 7401 7500: gap of unknown length
* 7501 13700: contig of 6200 bp in length
* 13701 13800: gap of unknown length
* 13801 16982: contig of 3182 bp in length
* 16983 17082: gap of unknown length
* 17083 22047: contig of 4965 bp in length
* 22048 22147: gap of unknown length
* 22148 27289: contig of 5142 bp in length
* 27290 27389: gap of unknown length
* 27390 41538: contig of 14149 bp in length
* 41539 41638: gap of unknown length
* 41639 54952: contig of 13314 bp in length
* 54953 55052: gap of unknown length
* 55053 81486: contig of 26434 bp in length
* 81487 81587: gap of unknown length
* 81587 107287: contig of 25701 bp in length.
FEATURES
Location/Qualifiers
1..107287
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-1019"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 9.5%; Score 136.4; DB 2; Length 107287;
Best Local Similarity 61.0%; Pred. No. 3.8e-10;
Matches 221; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
Qy 45 TCTATGATTTTCAGATTTGCATAGAATTTCTATCTATCAGAGACGCCCTGCAGAGATCC 104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84701 TCCTTCATAGCTTGTGCGGATTAACCTTATGTGCACTCTGCGACGCTTTCATCTTTCT 84642
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 105 CAAATTAGTCTAAAAATTAATCTTCAGTCTCGGAACCAACTCAGGACCCAAACCCGTCGC 164
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84641 CTTAATAGCTTAACCTTCTCTGTAGTCTGCTGAACAATATTCGTCCTCCCACTACACTC 84582
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 165 TCACCCCACTCAGTCTAATATACAGATATGACACTTATGACCATATAGAGCTCGTAA 224
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84581 TCACCCGACTCAAAACCAACATAACAGAGTTCTCGACCTCCGACCATACAAGCTTCAAT 84522
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 225 GGTGCCATCTAGATGCCAGATTGGAAACTGTTATTGTAGCGGAACCTCAACTAAGCGTAAA 284
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84521 GGTGCCATCTCTATACAGATAGTAATGTTGTTGTACGTGAACCTATCAACGGTAGA 84462
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 285 AAATCTCTCAACTACTCTTAGTAATAATACATAGCTCCAAATCGTATCTCTTAGTATA 344
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84461 TGACTATCCCAAGTCCCACTTGTATCAAGTACACAACTCTCAACAATCTCTTAACGAC 84402
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 345 TGAATCACCTTCTCAATTCACCATCGGTCTGAGGATGGAATGCAGACCGGTGCCACGA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84401 TGAATCGTCTCTCTGACTGACCATCTGTTTGGGATGATAGCCGAACCTCAACCTCAAC 84342
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 405 TT 406
|||
Db 84341 TT 84340
|||
Search completed: August 13, 2005, 11:00:00
Job time : 10343 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 02:09:43 ; Search time 7384 Seconds
(without alignments)

7387.066 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggacgcgctgttc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	224.4	15.7	581	9	CL864797 TM1-GSS00
2	183.8	12.8	768	9	CL251651 ZMMBBb060
3	175.4	12.2	758	9	CL404830 ZMMBBb039
4	172.8	12.1	1006	9	CL263790 ZMMBBb062
5	172	12.0	447	9	CL864019 TM1-GSS00
6	164.6	11.5	973	9	CL249359 ZMMBBb059
7	162.4	11.3	762	9	CL863878 TM1-GSS00
8	157.6	11.0	643	9	CL864239 TM1-GSS00
9	155.8	10.9	355	9	CL202255 ZMMBBb056
10	151.6	10.6	1057	9	CL285168 ZMMBBb062
11	147.4	10.3	957	9	CL291015 ZMMBBb063
12	146.4	10.2	739	9	CL864139 TM1-GSS00
13	143.2	10.0	569	8	BH021651 GH_MBB000
14	142	9.9	686	8	BH022375 GH_MBB000
15	139.8	9.8	725	9	CC746178 ZMMBBb012
16	139.8	9.8	781	8	BH022260 GH_MBB000
17	139.6	9.7	750	9	CC688457 OGWC337TH
18	139.6	9.7	1015	9	CL010422 ZMMBBb055
19	137.4	9.6	509	4	BI122809 I013P31P
20	137.4	9.6	799	8	BZ702320 PUCD182TD
21	137.2	9.6	660	8	BH021854 GH_MBB000
22	137	9.6	694	8	BZ420292 if66d04.b
23	136.8	9.5	697	8	BH985303 oeil6d03
24	136.8	9.5	843	9	CG006002 ZUACL81TV

25	136.4	9.5	679	9	CL243740
26	136.4	9.5	735	9	CC752373
27	136.4	9.5	786	9	CC757350 ZMMBBb014
28	136.4	9.5	893	9	CC979403 ZUABJ09TH
29	136.4	9.5	931	9	CC978577 ZUAGY07TH
30	136.4	9.5	957	9	CL248255 ZMMBBb059
31	135.8	9.5	750	9	CC783903 ZMMBBb014
32	135.8	9.5	792	9	CC986980 ZUAD810TH
33	135.8	9.5	818	9	CC815980 ZMMBBb052
34	135.8	9.5	848	9	CG245637 OGXB12TV
35	135.8	9.5	857	9	CG909847 ZMMBBb034
36	135.2	9.4	651	8	BZ488595 B0NCY26TF
37	135.2	9.4	653	9	CC962470 BOIGF89TR
38	135.2	9.4	708	9	CC760866 ZMMBBb015
39	135.2	9.4	797	8	BZ457826 BONLD73TF
40	135.2	9.4	805	8	BZ453701 BONFS55TF
41	134.8	9.4	629	8	BZ719317 PUCFH65TD
42	134.8	9.4	671	9	CG867526 ZMMBBb035
43	134.8	9.4	787	9	CC603679 OGWES48TV
44	134.6	9.4	951	9	CG300841 OGOC059TV
45	134.2	9.4	705	8	CC146564 ZMMBBb000

ALIGNMENTS

RESULT 1
LOCUS CL864797
DEFINITION TM1-GSS000073r BAC and BIBAC libraries from Upland cotton genetic standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV149b20 5', genomic survey sequence.

ACCESSION CL864797
VERSION CL864797.1
KEYWORDS GI:51321527
SOURCE GSS.

ORGANISM Gossypium hirsutum (upland cotton)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 581)
AUTHORS Xu,Z., Kohel,R.J., Zhang,H.B., Dong,J., Covalada,L., Lee,M., Koo,P. and Yu,J.Z.
TITLE Genome-Wide Synteny between Arabidopsis and Cotton
JOURNAL Unpublished (2004)
COMMENT Other GSSs: TM1-GSS000073f
Contact: John Z. Yu

Cotton Molecular Genetics Laboratory
USDA-ARS Crop Germplasm Research Unit
2765 F&B Road, Building 11, College Station, TX 77845. USA
Tel: 979-260-9237
Fax: 979-260-9333
Email: zyu@qutun.tamu.edu

for more detail, please see
http://algonon.tamu.edu/htdocs-cotton/cottondb.html

Seq primer: sp030
Class: BAC ends

High quality sequence stop: 581.

Location/Qualifiers

1. 581
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/cultivar="TM-1"
/db_xref="taxon:3635"
/clone="GH-TM1-CBV149b20"
/tissue_type="young leaves"
/lab_host="DH10B"
/clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
/note="vector: pCLD04541; pBeloBAC11; For more details on library construction, ordering clones and sequence analysis see
http://algonon.tamu.edu/htdocs-cotton/cottondb.html"

FEATURES
source

ORIGIN

Query Match 15.7%; Score 224.4; DB 9; Length 581;
Best Local Similarity 77.1%; Pred. No. 3.1e-27;
Matches 273; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 36 TGCTTGATATCATGATTTTCAGATTTGCATAAGACTTCTATCTATCATCAGAAAGCGCTGC 95
Db 195 TACTCAATATCATGCTTTTTCAGATTTGCATAAGACTTCTATCTATCATCAGAAAGCGCTGC 254

Qy 96 AGAGGATCCCAATAGTCTAAATATCTTCAAGTCTCGGAAACCACTCAGGACCCAAA 155
Db 255 AGACGATCTCAACACAGTTTAACTTGCCTTCAGTCTCAAAAACCAACTCAGGACCCAAA 314

Qy 156 ACCGTCGCTCACCAACTGATCTAATAAACAAGATGATGACACTTTATGACCATATAGA 215
Db 315 ACTTGATGCTTACCCTCACTGATCTAATAGTGAATGACACTTTACCAACCATATAA 374

Qy 216 GCCTCGTAAGTGCATCTAGATGCCAGATTGGAACCTGTTATTGTAGGCGAACTCAACT 275
Db 375 GCCTCGTAAGCGCATATGATGCTAGACTAGAACTGTTATTGTAGGCGAACTCAGCT 434

Qy 276 AACGGTAAAAATCCTCTCAACTACCTTAGTAATAATCATAGTCTCCAAATCGTATCC 335
Db 435 AGTGGCAGATAATCCTCCCACTACCTCAGAAATCAATCACAAATTTTGAAGTATATCC 494

Qy 336 TCTAGTATGAATCACTCTCAATTTGACCATCGGTCTGAGGATGGAATGCA 389
Db 495 TTTGGTATCTGAATCACCTCTTAGATTGACACCGGTCTGAGGATGGAACGCA 548

RESULT 2

CL251651/c
LOCUS ZM5B0604M16f ZM5Bbb (HindIII) Zea mays genomic clone
DEFINITION ZM5B0604M16 5', genomic survey sequence.

ACCESSION CL251651

VERSION CL251651.1 GI:41354780

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 768)

Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,

Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.

Sequencing of the maize genome at PGR (2003c)

Unpublished (2003)

Contact: Bharti,A.K.

Dr.Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers

University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 430.

Location/Qualifiers

source

1. .768

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZM5B0604M16"

/lab_host="E. coli DH10B"

/clone_lib="ZM5Bbb (HindIII)"

/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 12.8%; Score 183.8; DB 9; Length 768;

Best Local Similarity 74.8%; Pred. No. 1.5e-20;
Matches 270; Conservative 0; Mismatches 87; Indels 4; Gaps 3;

Qy 39 TTGATATCTATGATTTTTCAGATTTGCATAAGACTTCTAT-CTATCAGAAAGCGCTGCAG 97

Db 424 TAGATATCCCAATTTCTCAGATCTGAATAAGATATATATATATCAGAGCTGCCTTAAG 365

Qy 98 AGGATCCCAATTTAGTCTAAAAATTTCTTCAGTCTCGGAAACCACTCAGGAC-CCAAAA 156

Db 364 AGGATCCCAATTTAGTCTAAAAATTTCTTCAGTCTCGGAAACCACTCAGGAC-CCAAAA 305

Qy 157 CCGTCGCTCACCACCACTCAGTCTAATAAACAAGATGATGACACTTTATGACCATATAGAG 216

Db 304 TCTGTGCTCGCCCACTTAGTTTCAACAATAAGTACGACACTTACGACCATACAGAG 245

Qy 217 CCTCGTAAAGTGCCATCTAGATGCCAGATTGGAACCTGTTATTGTAGGCGAACTCAACTA 276

Db 244 CTTCTGAAGTGCTATTTTGAATCTAGACTGGAACCTATTTATTAAGTGAACCTCGACTA 185

Qy 277 ACGGTAAAAATCCTCTCAACTACCTTAGTAAATCAATCAATCAATCAATCAATCAATCAAT 336

Db 184 ACGGTAAAAATCCTCTCAACTACCTTAGTAAATCAATCAATCAATCAATCAATCAATCAAT 125

Qy 337 CTAGTATATGAATCACTCTTCT-CAATTTGACCATCGGTCTGAGGATGGAATCGACACCG 394

Db 124 CCAGTATCTAAATCACCTCTCCCTAAGAAAGATCTATCTAAGGATGGAATGTAGTACT 65

Qy 395 G 395

Db 64 G 64

RESULT 3

CL404830/c

LOCUS ZM5B0397006f ZM5Bbb (HindIII) Zea mays genomic clone

DEFINITION ZM5B0397006 5', genomic survey sequence.

ACCESSION CL404830

VERSION CL404830.1 GI:45216049

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 758)

Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,

Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.

Sequencing of the maize genome at PGR (2003c)

Unpublished (2003)

Contact: Bharti,A.K.

Dr.Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers

University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 89.

Location/Qualifiers

source

1. .758

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZM5B0397006"

/lab_host="E. coli DH10B"

/clone_lib="ZM5Bbb (HindIII)"

/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 12.2%; Score 175.4; DB 9; Length 758;
 Best Local Similarity 75.1%; Pred. No. 3.6e-19;
 Matches 259; Conservative 0; Mismatches 81; Indels 5; Gaps 3;

QY 52 TTTTCAGATTGCAATAGACTTCTATCTATCAGAGAGCGCTGCAGAGGATCCCAATTA 111
 |||||
 DB 750 TATTCAAATTTGCATATGA-TTTTGTCTATCAGAGCGCGTTTCCAAACGATTCGGAATCA 692
 |||||

QY 112 GTCTAAATATCTTTCAGTCTCGGAAACCACTCAGGACCCAAACCGCTCGCTCACCCA 171
 |||||
 DB 691 ATCTAACCTATCTCTATCTCAGAGACCAACTCTGGACCCAGAACAGCTCGCTCGCCCA 632
 |||||

QY 172 ACTCAGT-CTAATAAACAAGATGACACTTATGACCATATAGAGCTCTGTAAGTGCC 230
 |||||
 DB 631 ACTTAGTCCCAACGTAAGGAGTGCACACTTACAACCGTACAGTGCCTCGTAAGTGCC 572
 |||||

QY 231 ATCTAGATGCCAGATTGGAACTGTTATTTAGGCGAACTCACTACGGTAAAAATCC 290
 |||||
 DB 571 ATCTGGATGCTAGACTGGAAGCTATTATTCTAGGCAAACTCTGCTAA---TGCAAGTAC 515
 |||||

QY 291 TCTCAACTACCTTAGTAATAATCACATAGTCTCCAAATCGTATCTCTAGTATATGAATC 350
 |||||
 DB 514 TCCCACTACATCAGAAATCATAAATCACTCTCAACATGCTCCTCAGTATCTGAATC 455
 |||||

QY 351 ACCTTCTCAAATTCACCATCGCTGAGGATGGAATGCAGACCGG 395
 |||||
 DB 454 ATCTCTCTAACTGAGCGTCTGTTTGAGGATGAAACACAGTACTG 410
 |||||

RESULT 4
 CL263790 1006 bp DNA linear GSS 02-FEB-2004
 LOCUS ZMWB0623P14f ZMWB0623P14 5', genomic survey sequence.
 DEFINITION ZMWB0623P14 5', genomic survey sequence.

ACCESSION
 VERSION CL263790.1 GI:41627929
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1006)
 Bharti,A.K., Young,S., Kaychok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGR (2003c)
 Unpublished (2003)
 CONTACT: Bharti,A.K.
 Dr.Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu

Seq primer: T7
 Class: BAC ends
 High quality sequence start: 51.
 Location/Qualifiers
 1..1006
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMWB0623P14"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMWB0623P14 (HindIII)"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

FEATURES
 source
 1..1006
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMWB0623P14"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMWB0623P14 (HindIII)"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN
 Query Match 12.1%; Score 172.8; DB 9; Length 1006;
 Best Local Similarity 69.7%; Pred. No. 9.2e-19;
 Matches 248; Conservative 0; Mismatches 107; Indels 1; Gaps 1;

QY 36 TGCTTGATATCTATGATTTTTCAGATTGCAATAGACTTCTATCTATCAGAGAGCGCTGC 95
 |||||
 DB 96 TACTCAATCTCTTGGCATTTTAAGATCCGCGTAAGACTGCTACCTATCAGATGCTTCTCTC 155
 |||||

QY 96 AGAGATCCCAATTTAGTCTAAAATTTATCTTTCAGTCTCGGAAACCACTCAGGACCCAAA 155
 |||||
 DB 156 AGTTGGTCTCGAATCAGTTTACCTTTCTCTTCGGTATCAGAAAATAAATCTGCGCCCTAGA 215
 |||||

QY 156 ACCGTCGCTCACCAACTCAGTCTAATAATAACAGAGTATGACACTTATGACCATATAGA 215
 |||||
 DB 216 ACTCGTCTCGGCCAACTCAGTCCCAACAGTAGGATGACACATACACCATATAC 275
 |||||

QY 216 GCCTCGTAAAGTGCCATCTAGATGCCAGATTGGAACTGTTATTTAGGCGAACTCAACT 275
 |||||
 DB 276 GCCTCGTACGCTGCTATTTGAATACGCTCTGGTAGCTATTATTATAAGCGACTCTACT 335
 |||||

QY 276 AAGCGTAAAAATCTCTCACTACCTTAGTATTAATATCAC-ATAGCTCCAAATCGTATC 334
 |||||

DB 336 AGCGGCAATAGTCTTTCCAGGTGCGCTCGGAAATCCATTAATCTTATTTTCTTAACATATC 395
 |||||

QY 335 CTCTAGTATATGAATCACTTCTCAATTCACCATCGTCTGAGGATGGAATGCAG 390
 |||||

DB 396 CTCATATTTGAATCTCTCTTCTGATGACCATCTATCTGGGATGAACGCGAG 451
 |||||

CL864019 447 bp DNA linear GSS 19-AUG-2004
 LOCUS TM1-GSS000206f BAC and BIBAC libraries from Upland cotton genetic standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV047P09 5', genomic survey sequence.

ACCESSION
 VERSION CL864019.1 GI:51320749
 KEYWORDS GSS.
 SOURCE Gossypium hirsutum (upland cotton)
 ORGANISM Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 447)
 Xu,Z., Koneh,R.J., Zhang,H.B., Dong,J., Covalada,L., Lee,M., Koo,P. and Yu,J.Z.
 Genome-Wide Synteny between Arabidopsis and Cotton
 Unpublished (2004)
 Other GSSs: TM1-GSS000206r
 CONTACT: John Z. Yu
 Cotton Molecular Genetics Laboratory
 USDA-ARS Crop Germplasm Research Unit
 2765 F&B Road, Building 11, College Station, TX 77845. USA
 Tel: 979 -260-9237
 Fax: 979-260-9333
 Email: zyu@qutun.tamu.edu
 for more detail, please see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html

Seq primer: sp010
 Class: BAC ends
 High quality sequence stop: 447.
 Location/Qualifiers
 1..447
 /organism="Gossypium hirsutum"
 /mol_type="genomic DNA"
 /cultivar="TM-1"
 /db_xref="taxon:3635"
 /clone="GH-TM1-CBV047P09"
 /tissue_type="young leaves"
 /lab_host="DH10B"
 /clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
 /note="vector: pCLD04541; pBelobAC11; For more details on library construction, ordering clones and sequence analysis see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

FEATURES
 source
 1..447
 /organism="Gossypium hirsutum"
 /mol_type="genomic DNA"
 /cultivar="TM-1"
 /db_xref="taxon:3635"
 /clone="GH-TM1-CBV047P09"
 /tissue_type="young leaves"
 /lab_host="DH10B"
 /clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
 /note="vector: pCLD04541; pBelobAC11; For more details on library construction, ordering clones and sequence analysis see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

ORIGIN

Query Match 12.0%; Score 172; DB 9; Length 447;
Best Local Similarity 67.7%; Pred. No. 1.4e-18;
Matches 241; Conservative 0; Mismatches 115; Indels

35	Qy	GTGCTTGATATCTATGATTTCTCAGATTTTGCATAAGACCTTCTATCTATCAGAGAGCGCTG	94
417	Db	GTATTCAATCTCTTACGGCTTCAGATCTGCATACGACTTTGGCCTTATCAGACGCTTCCTT	358
95	Qy	CAGAGATCCCAAAATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAA	154
357	Db	TAATCGATCTCTAAATTAACCTGACCTTATCTCTCAGTATCTGCCACCAACTCAAGTCCAAG	298
155	Qy	AACCCGTGCTCACCACCTGAGTCTTAATATAACAGAGTATGACACTTATGACCATATAG	214
297	Db	AACCTTGCTCTCCCGAGTTCAGTCCAAACACTAGGTGTTTGACACCTTCGCCCAACAG	238
215	Qy	AGCCTCGTAAGGTGCCATCTAGATGCCAGATTGGAAACTGTTTATTGTAGGCGAACTCAAC	274
237	Db	TGCTTCATAAGGTGCCATCCGAATACTCGACTGATAACTATTGTTGTATGCAAACTCTTC	178
275	Qy	TAAACGTAATAAAATCCTCTCAACTACCTTAGTAATAAATACATAGCTCCAAATCGTATC	334
177	Db	CAAAGCAAGTAGTCTCTCCAACTACTCCGGAATCAATAACACATCCCTCAGCATGTC	118
335	Qy	CTCTAGTATATGAATCAACCTTCTCAAATTGACCATCGGTCTGAGGATGGAATGCAG	390
117	Db	TTCCGAATCTCGAATAACTCTTTTCGACGTGACCATCAGTTTGGGATGGAAGCG	62

RESULT 6 :
CL249359/c

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
CL2493359/C	ZMMBBb00595C21f ZMMBBb (HindIII) Zea mays genomic clone ZMMBBb00595C21 5', genomic survey sequence.	CL2493359	CL2493359.1	GI:41105913	GSS.	Zea mays
		CL2493359				Zea mays
						Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 973)
REFERENCE	AUTHORS					Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE	JOURNAL					Sequencing of the maize genome at PGIR (2003c) Unpublished (2003)

FEATURES source

```

FEATURES
1..973
location/Qualifiers
    organism="Zea mays"
    mol_type="genomic DNA"
    cultivar="B73"
    db_xref="taxon:4577"
    clone="ZMMBB0595C21"
    lab_host="E. coli DH10B"
    clone_lib="ZMMBB (HindIII)"
    note="Vector: pCUI (HindIII)"
    note="Site 1: HindIII; Site 2: HindIII"

```

Query Match

Query Match 11.5%; Score 164.6; DB 9; Length 973;

	Best Local Similarity	65.8%;	Pred.No. 2.1e-17;		Mismatches	124;	Indels	0;	Gaps	0;
	Matches	239;	Conservative	0;						
Qy	28	CGGGCTGGTCTTGATATCTATGATTTTCAGATTTGCATAAGACTTCTATCTATCAGAAG	87							
Db	865	CGACTTGGAACTCTATATCTTTGCGTTTAAGATCTGCATACGATTTTGTGACAACTGGAAG	806							
Qy	88	ACGCGCTGCAGAGGATCCCAAAATTAGTCTAAAAATTATCTTCAGTCTCGGAACCAACTCAG	147							
Db	805	CTTCTTTTCAAACATATCTCGGATCACTTTCACCTTTTTCCTCAGTCTCGTGAACCAAAATCAA	746							
Qy	148	GACCCAAAACCCGTGCGTCAACCACTCAGTCTTAATATAACAGAGTATGACACTTATGAC	207							
Db	745	CCCGGCATATCTTTTCTCACTGAGCTCAGTCTAGTATATATAGAGTTGCGACATTCAGCGAC	686							
Qy	208	CATATAGAGCCTCGTAAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATTGTGAGCGGA	267							
Db	685	CATATAAGGTTTCATGTGGTGCCATCTTTATATCTCAACTGAAACATATTATTATACGTGA	626							
Qy	268	ACTCAACTAAGGTAAAAAATCTCTCAACTAGTAAATAAAACACATAGCTCCAAA	327							
Db	625	ATTTGACCAATGGTAAATATTTCTCCAGTTGCTTTAAACTCTAGTACACAACTTGAA	566							
Qy	328	TCGTATCCTCTAGTATATGAATCACTTCTCAAATTTGACCATCGGTCTGAGGATGGAATG	387							
Db	565	GCATATCTTCGAGAAATTTGAATCACAGCTCGGACTGACCCTCAGTCTGAGGGGTGAAAG	506							
Qy	388	CAG 390								
Db	505	TAG 503								

RESULT 7

CL863878	CL863878	762 bp	DNA	linear	GSS 19-AUG-2004			
LOCUS	TM1-GSS000151f BAC and BIBAC libraries from Upland cotton genetic standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV0260U5', genomic survey sequence.							
DEFINITION	genomic survey sequence.							
ACCESSION	CL863878							
VERSION	CL863878.1	GI:51320608						
KEYWORDS	GSS.							
SOURCE	Gossypium hirsutum (upland cotton)							
ORGANISM	Gossypium hirsutum							
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.							
REFERENCE	1 (bases 1 to 762)							
AUTHORS	Xu, Z., Kohel, R.J., Zhang, H.B., Dong, J., Covaleda, L., Lee, M., Koo, P. and Yu, J. Z.							
TITLE	Genome-Wide Synteny between Arabidopsis and Cotton							
JOURNAL	Unpublished (2004)							
COMMENT	Other GSSs: TM1-GSS000151r							

FEATURES

```

1. .762
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/cultivar="TM-1"
/db_xref="taxon:3635"
/clone="GH-TM1-CBV026005"
/tissue_type="Young leaves"
/lab_host="DH10B"

```

/clone lib="BAC and BIBAC libraries from Upland cotton
genetic standard TM-1"
/note="vector: pCLD04541; pBelOBAC11; For more details on
library construction, ordering clones and sequence
analysis see
http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

ORIGIN

Query Match 11.3%; Score 162.4; DB 9; Length 762;
Best Local Similarity 65.4%; Pred. No. 5e-17;
Matches 238; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
Qy 27 CCGGGCTGGCTGATATCTATGATTTTCAGATTTGCATAGAGCTTCTATCTATCAGAA 86
Db 153 CCCATCTGATCTCAATATCTCTTTGTTTCAATCCGCATACGACTTCTGAGCATCTGTG 212
Qy 87 GACGCTGCAGAGATCCCAATATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCACTCA 146
Db 213 GCTGCTTCAGACTTTCACGGATTACTTTTACTTTCTGTTCCAGCATCTTTAATCAAAATCA 272
Qy 147 GGACCCAAACCCGCTCGCTACCCCACTCAGCTCTAATATACAGATGACACTTATGA 206
Db 273 ACTCCAAAATTTTACTTTTACCGAGCTCGGTCCAAACAAATGGGTACAGCATTTGGA 332
Qy 207 CCATATAGAGCTCTGTAAGTGCCATCTAGATGCGAGATTGGAATCTTTATTGTAGCG 266
Db 333 CCGTACAAAGCTCTGTAAGTGCCATCTTATATCTTGATTGAAACTATTGTTGTAAGCA 392
Qy 267 AACTCAACTAACCGGTAAATAATCTCTCAACTAGCTTATAGTAAATCAATACATAGTCAA 326
Db 393 AATTCAATCAAAAGGTAAATACCGTTCCTCACTCAACCACTCAAGGATCGAGCATCTC 452
Qy 327 ATCGTATCTCTAGTATATGATCACTTCTCAATTCGACCATCGGTCTGAGGATGGAAT 386
Db 453 AACATATCTCAAGTATCTGAATTAATCCGCTCGGATTCACCATCGGTTTGGGGATGAAA 512
Qy 387 GCAG 390
Db 513 GCAG 516

RESULT 8
CL864239 643 bp DNA linear GSS 19-AUG-2004
LOCUS TM1-GSS000337f BAC and BIBAC libraries from Upland cotton genetic
DEFINITION standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV088C22 5',
genomic survey sequence.
ACCESSION CL864239
VERSION CL864239.1 GI:51320969
KEYWORDS GSS.
SOURCE Gossypium hirsutum (upland cotton)

Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 643)
AUTHORS Xu,Z., Kohel,R.J., Zhang,H.B., Dong,J., Covalada,L., Lee,M., Koo,P.
and Yu,J.Z.
TITLE Genome-Wide Synteny between Arabidopsis and Cotton
JOURNAL Unpublished (2004)
COMMENT Other GSSs: TM1-GSS000337r
Contact: John Z. Yu
Cotton Molecular Genetics Laboratory
USDA-ARS Crop Germplasm Research Unit
2765 F&B Road, Building 11, College Station, TX 77845, USA
Tel: 979-260-9237
Fax: 979-260-9333
Email: zyu@qutn.tamu.edu
for more detail, please see
http://algodon.tamu.edu/htdocs-cotton/cottondb.html
Seq primer: gp010
Class: BAC ends
High quality sequence stop: 643.

FEATURES
source

Location/Qualifiers
1..643
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/cultivar="TM-1"
/db_xref="taxon:3635"
/clone="GH-TM1-CBV088C22"
/tissue_type="young leaves"
/lab_host="DH10B"
/clone_lib="BAC and BIBAC libraries from Upland cotton
genetic standard TM-1"
/note="vector: pCLD04541; pBelOBAC11; For more details on
library construction, ordering clones and sequence
analysis see
http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

ORIGIN

Query Match 11.0%; Score 157.6; DB 9; Length 643;
Best Local Similarity 64.6%; Pred. No. 3.1e-16;
Matches 235; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
Qy 27 CCGGGCTGGCTGATATCTATGATTTTCAGATTTGCATAGAGCTTCTATCTATCAGAA 86
Db 146 CCCACCTGATCTTAATGTCTTTTCGTTTCAATCCGCGTACGACTTCTGACGATCTGTG 205
Qy 87 GACGCTGCAGAGATCCCAATATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCACTCA 146
Db 206 GCTGCTTCAGACTTTCACGGATTACTTTTACTTTCTGTTCCAGCATCTTTAATCAAAATCA 265
Qy 147 GGACCCAAACCCGCTCGCTACCCCACTCAGCTCTAATATACAGATGACACTTATGA 206
Db 266 ACTTCGAAAATTTTACTTTTCGCGAGCTCGGTCCAAACAAATGGGTACGCGCATTTACGA 325
Qy 207 CCATATAGAGCTCTGTAAGTGCCATCTAGATGCGAGATTGGAATCTTTATTGTAGCG 266
Db 326 CCGTACAAAGCTCTGTAAGTGCCATCTTATATCTTGATTGAAACTATTGTTGTAAGCG 385
Qy 267 AACTCAACTAACCGGTAAATAATCTCTCAACTAGCTTATAGTAAATCAATACATAGTCAA 326
Db 386 AATTCAATCAAAAGGTAAATACCGCTCCCATGACCACTCAAGGATCGAGCATCTC 445
Qy 327 ATCGTATCTCTAGTATATGATCACTTCTCAATTCGACCATCGGTCTGAGGATGGAAT 386
Db 446 AACATATCTCAAGTATCTGAATTAATCTGCTCGGATTCACCATCGGTTTGGGGGTGAAA 505
Qy 387 GCAG 390
Db 506 GCGG 509

RESULT 9
CL202255/c

LOCUS CL202255 355 bp DNA linear GSS 07-JAN-2004
DEFINITION ZMMBB0563L12r ZMMBBb (HindIII) Zea mays genomic clone
ZMMBB0563L12 3', genomic survey sequence.

ACCESSION CL202255
VERSION CL202255.1 GI:40718214
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 355)

AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE sequencing of the maize genome at PGR (2003c)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 180.
 Location/Qualifiers
 1. .355
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBBb063L12"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBBb (HindIII)"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

FEATURES

source

ORIGIN

Query Match 10.6%; Score 151.6; DB 9; Length 1057;
 Best Local Similarity 68.3%; Pred. No. 2.9e-15;
 Matches 239; Conservative 0; Mismatches 109; Indels 2; Gaps 2;
 QY 43 TATCTATGATTTTCAGATTGCGATAGACTTCTATC-TATCAGAACGCCCTGCAGGGA 101
 DB 78 TCTCTTTTATGTTTAAATACTGTATACGACCTGACTCAGGCGAGATCTCTTTTCAATCGG 137
 QY 102 TCCCAATTTAGTCTAAATAATTCTTTCAGTCTCGGAAACCAACTCAGGACCCAAACCCCGT 161
 DB 138 TCTCGAATCAGTTTACCTTATCTTCGGTATCAGAACTTAACCTCTGGGCTAGAACTCGC 197
 QY 162 CGCTCACCAACTCAGTCTTAATATAACAGAGTATGACACTTATGACCATATAGAGCTCG 221
 DB 198 CACTCGCTAGCTCAGTCCAAACAAATAGGAGTATGACACCTATGACCATATAATGCTCG 257
 QY 222 TAAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATTTAGCGGCACTCAACTACCGT 281
 DB 258 TAGGATGCCATTTCAATCTGGAATCTGCTAGCTATTAATATATACGCAAACTCTGCTAGCGC 317
 QY 282 AAAAAATCTCTCAACTACCTTAGTAATAAATCACATAGCTCCAAATCGTATCTCTTAGT 341
 DB 318 AATATATCTTCAATGCTTGGAAATCCATACACACCTCAACATATCTCCAGT 377
 QY 342 ATATGAATCACCTTCT-CAAAATTGACCATCGGTCTGAGGATGGAATGCGAG 390
 DB 378 ACTTGAATCACTTTTCCGATTGTCTATCTGCTGGGATGACAGCGAG 427

RESULT 10

CL285168

LOCUS

DEFINITION

CL285168 1057 bp DNA linear GSS 10-FEB-2004
 ZMMBBb0628K19r ZMMBBb (HindIII) Zea mays genomic clone
 ZMMBBb0628K19 3', genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1057)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGIR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr.Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 405.
 Location/Qualifiers
 1. .1057
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"

FEATURES

source

FEATURES

source

1. .957
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBBb063D02"
 /lab_host="E. coli DH10B"

/db_xref="taxon:4577"
 /clone="ZMMBBb0628K19"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBBb (HindIII)"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

CL291015 957 bp DNA linear GSS 10-FEB-2004
 ZMMBBb063D02r ZMMBBb (HindIII) Zea mays genomic clone
 ZMMBBb063D02 3', genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 957)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGIR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr.Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 399.
 Location/Qualifiers
 1. .957
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBBb063D02"
 /lab_host="E. coli DH10B"


```

/clone lib="Gossypium hirsutum L."
/note="vector: pCUIBAC-1; Site_1: HindIII; Site_2: NotI;
For more details on library preparation, ordering clones
and sequence analysis see
http://www.genome.clemson.edu/projects/stc/cotton/GH_MBB "

```

Query Match	10.0%	Score	143.2	DB	8	Length	569
Best Local Similarity	63.2%	Pred. No.	7.5e-14				
Matches	220	Conservative	0	Mismatches	128	Indels	0
Qy	46	CTATGATTTTCAGAGTTTGCATAAGACATCTCTATCTATCAGAGAGCGCTGCAGAGGATCCC	105				
Db	364	CTTTTCGGTTTCAAAATCGGCATACGATTTCTGACGATCTGTGGCTGCTTTTCAGACTTTTCAC	305				
Qy	106	AAATTAGTCTTAAAAATTATCTTCAGTCTCGGAACCAACTCAGGACCCCAAAACCGCTCGCT	165				
Db	304	GGATACCCCTTTACTTTCTGTTTCGGCACTTTTAATCAATCAATCCGAAATTTTACTTTT	245				
Qy	166	CACCCAACTCAGTCTTAATATAACAGAGATGACACTTATGACCATATAGACGCTCGTAAG	225				
Db	244	CACCGAGCTCGGTCCAAAACAATGGTGTA CGGCATTTACGCCGTCAAAAGGCTTCGTAAG	185				
Qy	226	GTGCCATCTAGATGCCAGATTGGAAA CTGTTTATTGTAGGCCGAATCAACTCAACGGTAAAA	285				
Db	184	GTGCCATCTTAAATACTTGATTGAAACACTATTGTGTAAAGCGAAATTCATCAAAAGGTAAT	125				
Qy	286	AATCCTCTCAACTACCTTTAGTAAATAAATACATAGTCCAAATCGTATCTCTAGTATAT	345				
Db	124	ACGGTTCCTATGAACCTACTGAATCTCGAGGATGCAACATCTCAACATATCCTCAAGTATCT	65				
Qy	346	GAATCACTCTTCAAAATTGACCATCGGTCTGAGGATGGAATGCGAGAC	393				
Db	64	GAATTAATCCGCTCGGATTGACCATCGGTTCGGGAAGAAAAGCAGAAC	17				

RESULT 14	BH022375/c	BH022375	686 bp	DNA	linear	GSS 29-JUN-2001
LOCUS		GH_MBB0002004r				
DEFINITION		GH_MBB0002004r, genomic clone GH_MBB0002004r, genomic survey sequence.				

FEATURES
SOURCE

RESULT 15
CC746178/C

LOCUS	CC746178	725 bp	DNA	linear	GSS 25-JUN-2003
DEFINITION	ZMMBBB0122010.f	ZMMBBB	zea mays	genomic clone ZMMBBB0122010 5', genomic survey sequence.	

SOURCE
ORGANISATION

REFERENCE AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 2.
2. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1.
3. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2.
4. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3.
5. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4.
6. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5.
7. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6.
8. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7.
9. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7, 8.
10. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also 1, 2, 3, 4, 5, 6, 7, 8, 9.

University of Arizona
Biological Sciences West, 488A, P.O. Box 210088, Tucson, AZ
85721-0088 USA

Class: BAC ends.
Location/Qualifiers
1. .725
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0122010"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. ssp. mays"

ORIGIN
Query Match 9.8%; Score 139.8; DB 9; Length 725;
Best Local Similarity 68.1%; Pred. No. 2.7e-13;
Matches 209; Conservative 0; Mismatches 97; Indels 1; Gaps 1;
Qy 80 ATCAGAAGACGCTGCAGAGATCCCAAATTAGTCTAAATTTATCTTCAGTCTCGAAAC 139
Db 308 ATCCGAGCTGCTTTCAGAGTATAACTTATCATTTTAACCTTTCTTCAGTTTGTGAT 249
Qy 140 CAATCTAGGACCCAAACCCGCTCGCTACCCAACTCAGTCTTAATATAACAGAGTATGACA 199
Db 248 CAAATCAACCCGAATACCTTTTCTTCCTGAGCTCAGTCTATTATAACAAAGTTTCGACA 189
Qy 200 CTTATGACCATATAGAGCTCGTAAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATT 259
Db 188 CTTCTGACCATACATGCTTCATACATTTGTCATTTTAAGCTCGATTGATTAACCTGTTTT 129
Qy 260 GTAGCGCAACTCAACTACGCTAAATAATCTCTCACTACCTTAGTAAATAAATCACATA 319
Db 128 ATATGCGAATTCAACTAATGTAGGAATCTCTCCCACTACCTTAAATTTCTAAACACA 69
Qy 320 GCTCAAATCGTATCCTCTAGTATATGAATCA-CCTTCTCAATTTGACCATCGGCTGAG 378
Db 68 ATGTCGAAGCATATCTTCAAGTATTTGAATCATAGGATTCAGATTGACCATCTGTCTGGG 9
Qy 379 GATGGA 385
Db 8 GATGAA 2

Search completed: August 13, 2005, 13:05:14
Job time : 7392 secs

THIS PAGE BLANK (1)